Microbial sterol side chain degradation in Actinobacteria

Wilbrink, Maarten Hotse

Microbial sterol side chain degradation in *Actinobacteria*

Cover picture: Jessica Wilbrink

The work described in this thesis was carried out in the Microbial Physiology group of the Groningen Biomolecular Sciences and Biotechnology Institute (GBB) at the University of Groningen, The Netherlands. The research was financially supported by a grant from the Integration of Biosynthesis and Organic Synthesis (IBOS) programme of the Advanced Chemical Technologies for Sustainability (ACTS) partnership of the Netherlands Organisation for Scientific Research (NWO). The author expresses his gratitude to Organon (currently part of MSD) and Solvay Pharmaceuticals (currently part of Abbott) for their financial support.

The author gratefully acknowledges financial support by the University of Groningen and the GBB for printing of this thesis.

Printed by Ipskamp Drukkers B.V., Enschede

ISBN: 978-90-367-4808-7 (printed version)
ISBN: 978-90-367-4807-0 (electronic version)

rijksuniversiteit groningen

Microbial sterol side chain degradation in *Actinobacteria*

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 11 maart 2011
om 13:15 uur

door

Maarten Hotse Wilbrink
geboren op 13 februari 1980
te Sappemeer

Promotor:	Prof. dr. L. Dijkhuizen
Copromotor:	Dr. R. van der Geize
Beoordelingscommissie:	Prof. dr. A.J.M. Driessen
	Prof. dr. D.B. Janssen
	Prof. dr. M.J. Larkin

CONTENTS

CHAPTER 1

General introduction

General introduction

THE ORDER OF THE *ACTINOMYCETALES*

The order of *Actinomycetales*, belonging to the *Actinobacteria*, comprises a wide variety of Gram-positive, aerobic bacteria with high GC (>60%) DNA. The order contains many suborders, including the Actinomycineae, Corynebacterineae, Frankineae, Propionibacterineae and Streptomycineae (Brands, 2005). Many of its members are soil dwelling, but they also inhabit aquatic and marine environments and plant or animal cells. *Actinobacteria* play a major role in the degradation of some of the most abundant organic compounds in Nature, such as the polymers cellulose and chitin, and are hence of great importance in the global carbon cycle (Lynd *et al*., 2002; Metcalfe *et al*., 2002). Furthermore, their ability to degrade an astonishing array of organic and xenobiotic compounds makes them invaluable in the production and modification of industrially relevant compounds as well as in bioremediation of soil and ground water (Seto *et al*., 1995; Pizzul *et al*., 2007).

Some of the actinobacterial genera are renowned for their ability to produce a wide variety of secondary metabolites, including the medically relevant antibiotics. Members of the genus *Streptomyces* are the most potent antibiotic producers known, accounting for an estimated two thirds of the clinically used antibiotics (Hopwood, 2007). Besides streptomycetes, also members of other genera produce antibiotics and compounds that are used in clinical medicine. For example, *Salinispora tropica* produces the anticancer drug Salinosporamide A, a proteasome inhibitor that has passed pre-clinical studies and human phase I trials (Adams, 2002; Feling *et al*., 2003). Also the diabetic drug acarbose, an α-glucosidase inhibitor derived from strains of *Actinoplanes* (Schmidt *et al*., 1977), is widely used in patients with type II diabetes to control blood glucose levels (Sim *et al*., 2010). Other well-known examples include the glycopeptide antibiotics vancomycin and teicoplanin that are produced by *Amycolatopsis orientalis* (Geraci *et al*., 1956) and *Actinoplanes teichomyceticus* (Bardone *et al*., 1978), respectively. These antibiotics are used as a last resort to treat infections of (multidrug resistant) Gram-positive bacteria like methycillin-resistant *Staphylococcus aureus* (MRSA) and enterococci (Wood, 1996). However, due to the huge increase in extensively drug-resistant bacteria, these drugs today are less efficacious in combating bacterial infections, thus stressing the importance of discovery and development of new antibacterial drugs (Srinivasan, 2002; Nailor and Sobel, 2009). Currently, efforts are made to develop modified versions of existing antibiotics and to discover new types of antibiotics on a global scale, e.g. by mining of actinobacterial genomes or metagenomes (Gottelt *et al*., 2010; Medema *et al*., 2010). This area of drug discovery has been extensively reviewed (Lorenz and Eck, 2005; Bode and Müller, 2005; Challis, 2008; van Wezel *et al*., 2009).

The relevance of *Actinobacteria* is reflected by the large number of actinobacterial genome sequences that are currently (2010, November) available in public databases; 163 out of a total of 1,371 prokaryotic genome sequences available in the NIH/NCBI public database are of actinobacterial origin. These include some of the most notorious human pathogens, like *Mycobacterium tuberculosis* and *Mycobacterium leprae*, the causative agents of tuberculosis (TB) and leprosy, respectively. *M. tuberculosis* annually claims an estimated one to two million lives and is the foremost bacterial killer worldwide (Frieden *et al*., 2003; WHO, 2005). It is estimated that about one third of the world population has a latent TB infection (WHO, 2005). In about 5-10% of the cases the infection progresses into active disease. TB is widespread in developing countries and is especially lethal to immunocompromised individuals, like HIV infected individuals (Frieden *et al*., 2003; WHO, 2005). The emergence of multidrug resistant (MDR) or extensively drug resistant (XDR) tuberculosis strains is a growing problem in Western countries as well, since many of the classic antibiotics previously used have lost their efficacy to combat these strains and thus these strains pose a serious threat to public health (Shah *et al*., 2007).

Biotechnological importance of microbial sterol side chain degradation

The global market for steroid based drugs and medicines has been estimated to be US$ 4-8 billion, while the market for 4-androstene-3,17-dione (AD) and 1,4-androstadiene-3,17-dione (ADD) alone is approaching US$ 1 billion (www.drugdiscoveryonline.com). Until some decades ago, the pharmaceutical industry solely relied on diosgenin (Fig. 1) derived from the Mexican yam (*Dioscorea composita*) for the production of steroids. Later, also solasodin was used as a raw material in the synthesis of steroid drug precursors (Kieslich, 1985). Over the years, diosgenin prices increased dramatically, mainly due to protection and thus limitation of the availability of *Dioscorea* plants. Therefore, the pharmaceutical industry started searching for alternative and cheap raw materials for the production of steroids.

Plant sterols like stigmasterol, β-sitosterol and campesterol were found to be relatively plentiful in soy bean oil, rape seed and paper pulp industry waste, providing suitable substrates for the chemical or microbial conversion into various steroids (Kieslich, 1985; Dias *et al*., 2002). Stigmasterol contains a C22-23 double bond, which enables the oxidative cleavage into progesterone, while mixtures of phytosterols can be degraded into C17-keto steroids by various patented *Mycobacterium* strains (Marsheck *et al*., 1972; Martin, 1977; Kieslich, 1985; Dias *et al*., 2002). Annually, more than 1000 tons of steroids are produced by microbial transformations, which can be further modified chemically and/or by bioconversion to form bioactive steroids with diverse functions (Schmid *et al*., 2001). *Mycobacterium* sp. NRRL B-3805 and a derived mutant were shown to produce the bioactive steroid testosterone from cholesterol and phytosterols, respectively (Liu and Lo, 1997; Lo *et al*., 2002). Other sources of steroid-like compounds are bile acids derived from cattle gall bladders, cholesterol from wool-grease and ergosterol produced by yeast which is used as substrate for the conversion into vitamin D and derivatives thereof (Kieslich, 1985; Dovbnya *et al*., 2010).

Sterol transforming Actinobacteria

An impressive number of *Actinobacteria* is able to degrade various sterol and steroid compounds (Szentirmai, 1990; Mahato and Garai, 1997; Fernandes *et al*., 2003). Table 1 lists a number of strains that are able to transform the side chains of various sterols, either as natural/UV mutants blocked in steroid ring degradation or requiring inhibitors of steroid ring degradation enzymes. Some of these strains are or have been used to produce steroid intermediates from cheap sterol sources at industrial- or laboratory scale. At Schering, Germany, the transformation of phytosterols from natural plant resources by *Mycobacterium* sp. mutants has been reported to account for the production of 200 tons of AD and ADD annually (Schmid *et al*., 2001).

Table 1. List of actinobacterial strains and/or mutants that are capable of sterol side chain degradation, either in the presence or absence of inhibitors of steroid ring degrading enzymes, thereby forming various steroid products. Adapted from Mahato and Garai (1997) and Fernandes *et al*. (2003).

Substrate	Microorganism	Product(s)	References
Cholesterol	*Mycobacterium phlei*	4-androstene-3,17-dione	Stadtman *et al*., 1954
Cholesterol	*Mycobacterium* sp. NRRL B-3805	4-androstene-3,17-dione	Liu and Lee, 1992
Cholesterol	*Rhodococcus corallina*	1,4-androstadiene-3,17-dione	Shi *et al*., 1992
Cholesterol	*Arthrobacter simplex and Mycobacterium* sp. NRRL B-3683	1,4-androstadiene-3,17-dione	Lee *et al*., 1993
Cholesterol	*Rhodococcus equi*	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione	Ahmed *et al*., 1993b
Cholesterol	*Mycobacterium* sp.	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione (iii) 17β-hydroxyandrosta-l,4-dien-3-one	Smith *et al*., 1993
Cholesterol	*Mycobacterium fortuitum* NRRL B-8153	3β-hydroxy-5-androstene-17-one	Srivastava and Patil, 1994
Cholesterol	*Mycobacterium* sp.	17β-hydroxy4-androstene-3-one	Liu *et al*., 1994
Cholesterol	*Rhodococcus equi*	(i) 1,4-androstadiene-3,17-dione (ii) 4-androstene-3,17-dione	Ahmed and Johri, 1993a
Cholesterol	*Rhodococcus rhodochrous* DSM43269 mutant RG32	(i) 1,4-androstadiene-3,17-dione (ii) 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid	(Wilbrink *et al*., submitted)
2α,3α-dihydroxy-5α-cholestan-6-one	*Mycobacterium vaccae*	(i) 2α,3α,6α-trihydroxy-5α-androstane-17-one (ii) 2α-hydroxyandrost-4-ene-3,17-dione	Vorbrodt *et al*., 1991
ergosterol ergosterol-3-acetate	*Mycobacterium* sp. VKM Ac-1815D	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione (iii) hydroxymethyl-4-pregnene-3-one	Dovbnya *et al*., 2010
ergosterol	*Mycobacterium* sp. NRRL B-3805	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione	Ambrus *et al*., 1995
19-hydroxy-cholesterol	*Rhodococcus* mutant k-3	(i) estra-l,3,5(10)-triene-3-ol (ii) 2(3-hydroxy-1,3,5(10)-estra-triene-	Murohisa and Iida, 1993a

		17-yl)-propionic acid (iii) 2-methyl-6(3-hydroxy-1,3,5(10)-estratriene-17-yl)-heptanoic acid (iv) 2(3-hydroxy-1,3,5(10),17-estra-tetraene-17-yl)-propionic acid	
19-hydroxy-campesterol	*Rhodococcus* mutant k-3	(i) 2(3-hydroxy-1,3,5(10),17-estra-tetraene-17-yl)-propionic acid (ii) 2,3-dimethyl-6-(3-hydroxy-1,3,5(10)-estratriene-17-yl)-heptanoic acid	Murohisa and Iida, 1993a
lanosta-7,9(11)-dien-3β-ol	*Mycobacterium* sp. NRRL B-3805	4,8(14)-androstadiene-3,17-dione	Wang *et al*., 1995
lithocholic acid	*Mycobacterium* sp.	20α-hydroxy-4-methylpregnene-3-one	Wang *et al*., 1994
3β-methoxy-ergosta-5,7,22-triene	*Mycobacterium*	3β-methoxymethoxy-21-hydroxy-20-methyl-5,7-pregnadiene	Weber *et al*., 1992
β-sitosterol	*Mycobacterium fortuitum*	9α-hydroxy-4-androstene-3,17-dione	Birke *et al*., 1992
β-sitosterol	*Arthrobacter oxydans*	(i) 3-oxo-4-cholene-24-oic acid (ii) 27-nor-4-cholestene-3,24-dione	Dutta *et al*., 1992
β-sitosterol	*Rhodococcus equi* k-3	(i) 3-oxo-1,4-ergostadiene-26-oic acid (ii) 3-oxo-4-ergostene-26-oic acid (iii) 20-carboxy-4-pregnene-3-one (iv) 20-carboxy-1,4-pregnadiene-3-one (v) 4-androstene-3,17-dione (vi) 1,4-androstadiene-3,17-dione	Murohisa and Iida, 1993b
β-sitosterol	*Mycobacterium* sp.	9α-hydroxy-4-androstene-3,17-dione	Borman *et al*., 1992
β-sitosterol	*Mycobacterium* NRRL B-3683	1,4-androstadiene-3,17-dione	Roy *et al*., 1992
β-sitosterol	*Arthrobacter simplex*	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione	Mathur *et al*., 1992
β-sitosterol	*Mycobacterium*	4-androstene-3,17-dione	Kurakov *et al*., 1993
β-sitosterol	*Mycobacterium vaccae*	4-androstene-3,17-dione	Spassov *et al*., 1993
β-sitosterol	*Nocardia* sp. M 29	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione	Martin and Wagner, 1976
β-sitosterol	*R. rhodochrous* DSM43269 mutant RG32	(i) 1,4-androstadiene-3,17-dione (ii) 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid	(Wilbrink *et al*., submitted)
solasodiene	*Mycobacterium* sp. NRRL B-3805	4-androstene-3,17-dione	Shukla *et al*., 1992
sterol	*Mycobacterium vaccae*	1,4-androstadiene-3,17-dione	Reiche *et al*., 1992
sterol	*Mycobacterium fortuitum*	9α-hydroxy-4-androstene-3,17-dione	Seidel and Hoerhold, 1992
sterol	*Mycobacterium* NRRL B-3805	4-androstene-3,17-dione	Lee, 1990
sterol	*Mycobacterium fortuitum*	9α-hydroxy-4-androstene-3,17-dione	Atrat *et al*., 1992
sterol	*Mycobacterium vaccae*	1,4-androstadiene-3,17-dione	Gottschaldt *et al*., 1993
sterol	*Arthrobacter simplex*	1,4-androstadiene-3,17-dione	Oda, 1994
sterol	*Mycobacterium* sp.	(i) 4-androstene-3,17-dione (ii) 1,4-androstadiene-3,17-dione	Zhang *et al*., 1992

THE GENUS *RHODOCOCCUS*

The genus of *Rhodococcus* belongs to the suprageneric group of mycolate containing nocardioform *Actinobacteria*, which includes among others the closely related genera *Mycobacterium*, *Nocardia*, and *Corynebacterium* (Goodfellow *et al*., 1998). Currently, the genus comprises 12 species that are validly described (Bell *et al*., 1998), while the total number of species within the genus may even exceed 40 (DSMZ strain collection). Their cell envelope comprises, in addition to the cytoplasmic membrane and the type IV peptidoglycan cell wall, another membrane layer consisting of mycolic acids that are linked to the arabinogalactan part of the cell wall matrix (Daffe *et al*., 1993; Sutcliffe, 1998). The rhodococcal mycolic acids are fatty acids that vary in length between 28 up to 54 carbon atoms and can contain multiple unsaturated bonds (Sutcliffe, 1998). The presence of mycolic acids in *Rhodococcus* spp. makes it difficult to classify them using classical Gram-staining and renders them partially or weakly acid-fast (Murohashi *et al*., 1969).
Generally, rhodococci are regarded as non-pathogenic, however, two exceptions are currently known. *Rhodococcus equi* is pathogenic to horses, foals in particular, and causes "rattles", a lung disease similar to tuberculosis in man. *Rhodococcus fascians* is a plant pathogen that causes, among others, leafy galls (LeChevalier, 1989). Interestingly, in both *R. equi* and *R. fasciens* the presence of virulence plasmids (the 80-90 kb plasmids p33701/p103 and the large linear plasmid pFiD188, respectively) are decisive factors for pathogenicity (Takai *et al*., 1991, 1993; Crespi *et al*., 1992).

Rhodococci are greatly appreciated for their ability to degrade various natural hydrocarbons as well as xenobiotic compounds, like phenols and polychlorinated biphenyls (PCBs) (Masai *et al*., 1995; Seto *et al*., 1995; Dabbs, 1998; Vaillancourt *et al*., 2003). Due to the importance of rhodococci in biodegradation and biotechnology, the topic has been reviewed extensively (e.g. van der Geize and Dijkhuizen, 2004; de Carvalho and da Fonseca, 2005; Larkin *et al*., 2005; Martínková *et al*., 2009). Several factors contribute to the successful use of rhodococci in biotechnology. Firstly, rhodococci are generally able to tolerate and even thrive in the presence of organic solvents, e.g. liquid benzene (Paje *et al*., 1997), possibly due to the presence of their mycolic acids, which enable them to adhere to oil/water interfaces (Neu, 1996). Furthermore, they are efficient in the uptake of hydrophobic compounds, which is facilitated by the secretion of surfactants (Bicca *et al*., 1999). Also, the huge set of catabolic enzymes encoded by rhodococcal genomes, including many oxygenases, oxidoreductases and β-oxidation enzymes makes them useful for industrial applications (Larkin *et al*., 2005; McLeod *et al*., 2006). Adequate systems for the genetic manipulation of rhodococci are nowadays available, enabling the construction of targeted single gene disruptions and even multiple (unmarked) gene deletion mutants (Masai *et al*., 1995; van der Geize *et al*., 2001, 2008a). Also, the relatively fast growth rate of most rhodococci, as compared to for instance mycobacteria, is advantageous for their use in industrial processes (McLeod *et al*., 2006).

The large-scale production of acrylamide using a nitrilase-overproducing *Rhodococcus* strain is exemplary for the successful use of rhodococci in an industrial process (Yamada and Nagasawa, 1994; Banerjee *et al*., 2002). Rhodococci are also used in the *in situ* bioremediation of contaminated soils. Furthermore, *Rhodococcus* strains are used in the conversion of indene into cis-(1S, 2R)-indandiol, an intermediate in the synthesis of the HIV protease inhibitor indinavir sulphate, which is used in the treatment of AIDS (Buckland *et al*., 1999). The use of naturally occurring or genetically engineered *Rhodococcus* strains in biodesulphurization (BDS) of fossil fuels and other recalcitrant compounds is another promising future prospect. It was found, for example, that the extremely resistant dibenzothiopene can be degraded *in situ* by *Rhodococcus* spp. (Akhtar *et al*., 2009). Currently, BDS is not economically feasible, but in view of recent technological advances in the field, its application may be realized in the near future (Kilbane, 2006; Martínková *et al*., 2009).

Lastly, rhodococci find applications in steroid drug synthesis. Many rhodococcal strains are excellent sterol and steroid degraders and generally well-suited for large scale fermentation of these compounds (Finnerty, 1992; van der Geize and Dijkhuizen, 2004; Larkin *et al*., 2005). During degradation of sterols, intermediates are produced that are suitable starting materials to synthesize bioactive steroids. Several studies have successfully identified genes involved in steroid and sterol degradation, and characterized the enzymes involved in these processes in various *Rhodococcus* spp. (Itagaki *et al*., 1990b; van der Geize *et al*., 2000, 2001, 2002a, 2002b, 2007, [Chapter 2], 2008b; Knol *et al*., 2008; Petrusma *et al*., 2009; Rosłoniec *et al*., 2009, [Chapter 3]; Mathieu *et al*., 2010). *Rhodococcus* mutants with multiple gene deletions in key steps of steroid nucleus degradation accumulate steroid pathway intermediates from cheap sterol precursors (van der Geize *et al*., 2003, 2009). The availability of the genome sequence of *Rhodococcus jostii* RHA1 in 2006, together with transcriptomic analysis of cholesterol grown RHA1 cells have resulted in deeper insight into the processes and genes involved in steroid and sterol degradation (McLeod *et al*., 2006; van der Geize *et al*., 2007, [Chapter 2]).

RHODOCOCCAL GENOME SEQUENCES

R. jostii RHA1 was isolated from lindane contaminated soil and is a potent polychlorinated biphenyl degrader (Seto *et al*., 1995; Masai *et al*., 1995). The 9.7 Mb genome sequence of strain RHA1 (GenBank accession number CP000431) was the first of the *Rhodococcus* genus and proved to be one of the largest bacterial genomes that is presently known (McLeod *et al*., 2006). The complete genome sequence of strain RHA1 indicated a linear chromosome composed of 7.9 Mb and three linear megaplasmids, pRHL1 (1.1 Mb), pRHL2 (440 kb) and pRHL3 (330 kb), with an average GC content of 67%. Bioinformatic analysis revealed that the strain RHA1 genome encodes an astonishing number of catabolic enzymes, including more than one thousand putative oxidoreductases, some 30 putative cytochrome P450 monooxygenases, and a large array of predicted β-oxidation enzymes (e.g. 72 CoA ligases and 55 thiolases) and was hence denoted a “catabolic powerhouse” (McLeod *et al*., 2006).

Also, the presence of nonribosomal peptide synthases (NRPS) and polyketide synthases (PKS) demonstrate that strain RHA1 may harbor an extensive secondary metabolism. Many rhodococci harbor megaplasmids that usually contain genes coding for catabolic enzymes that enable their host to grow on specific (chlorinated) hydrocarbons or other organic compounds. Examples hereof are the abovementioned plasmids of RHA1, p1CP from *Rhodococcus opacus* 1CP (740 kb) (König *et al*., 2004), pBD2 from *Rhodococcus erythropolis* BD1 (Stecker *et al*., 2003) and pDK1 (380 kb), pDK2 (330 kb) and pDK3 (750 kb) from *Rhodococcus* sp. strain DK17 (Kim *et al*., 2002; Choi *et al*., 2005).

Recently, the 5.1 Mb genome sequence of the horse pathogen *R. equi* 103S (http://www.sanger.ac.uk/Projects/R_equi/) became publicly available, followed by those of *R. erythropolis* PR4 (6.9 Mb, GenBank AP008957) and *R. opacus* B4 (7.9 Mb, GenBank AP011115). Additionally, several *Rhodococcus* genomes are currently being sequenced and the partial genome sequences of various strains, including *R. equi* ATCC 33707, *R. erythropolis* DSM8424, *R. opacus* PD630, *Rhodococcus aetherivorans* I24 and *Rhodococcus rhodochrous* are known. The availability of these genome sequences emphasizes the importance of the genus *Rhodococcus* and contributes to a better understanding of the metabolic capabilities of its members. Currently, there is an extensive molecular toolbox available for *Rhodococcus* that enables the construction of targeted gene deletion or disruption mutants, as well as inducible expression systems for over-expression of heterologous proteins (van der Geize *et al*., 2000; 2001; 2008b; Nakashima and Tamura, 2004). These developments facilitate the further use of *Rhodococcus* strains and their enzymes in various fields of biotechnology.

STEROIDS

Steroids are naturally occurring hydrophobic molecules that have the perhydro-1,2-cyclopentanophenantrane ring system in common (Fig. 1). They are comprised of three six-membered rings A, B and C, and one five membered ring denoted D. Steroids form a large class of compounds with regulatory functions in eukaryotes or act as important constituents of eukaryotic membranes. Generally, steroids (and sterols, see below) are absent in

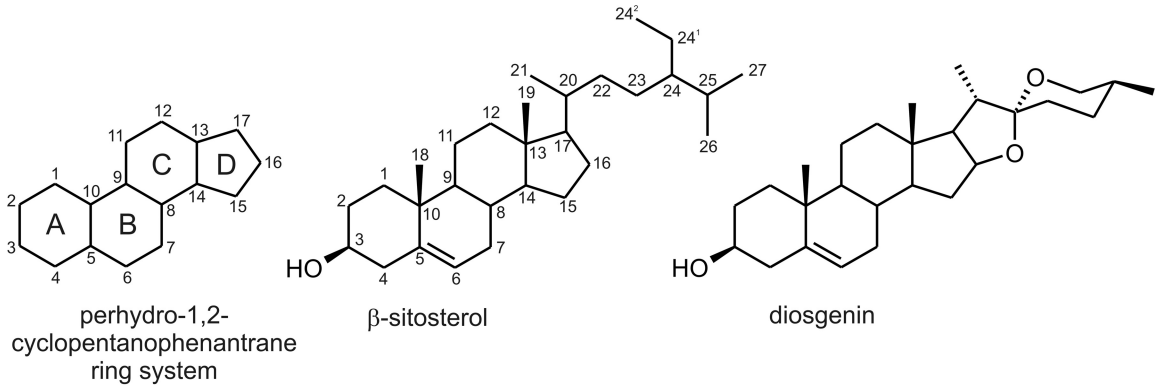


Fig. 1. Representation of the perhydro-1,2-cyclopentanophenantrane ring system with the rings denoted A-D and carbon numbering. β-sitosterol, the most abundant plant sterol and the plant steroidal sapogenin diosgenin.

bacteria, but they can be degraded by some bacteria and used as carbon and energy source for growth. In animals, steroids act as hormones at low (nanomolar range) concentrations. Biologically active steroids exist in many variations. The most important modifications are the degree of saturation of the ring system, the substituents present, mainly at carbons 3 and 17, or the presence of an aromatic A-ring as found in the estrogens. Steroids have important regulatory roles in reproduction, carbohydrate and protein catabolism, stress and immune responses and behavior. The former are important in regulating sexual processes and are mainly produced in the reproductive organs. The most important sex steroids are testosterone, estrone, estradiol and progesterone. The adrenal cortical steroids are produced in the adrenal cortex and can be divided in glucocorticoids and mineralocorticoids and are involved in e.g. carbohydrate and protein catabolism, stress responses, immune response and behavior. Also in plants, the polyhydroxylated brassinosteroids were found to have a signaling function (Vert *et al*., 2005; Gendron and Wang, 2007).

STEROLS

Sterols are hydrophobic molecules consisting of the aforementioned steroid ring system with an aliphatic side-chain at carbon 17; they generally possess a (β-)hydroxyl group at carbon three and are hence also named steroid alcohols (Fig. 1). Sterols such as cholesterol (Fig. 2) have an important structural role in determining membrane fluidity in all eukaryotes. The polar hydroxyl group of the sterol molecule is aligned with the polar phosphate group of phospholipids, while the hydrophobic ring and side-chain are positioned between the fatty acid tails, rendering the membrane more rigid and less permeable to small molecules

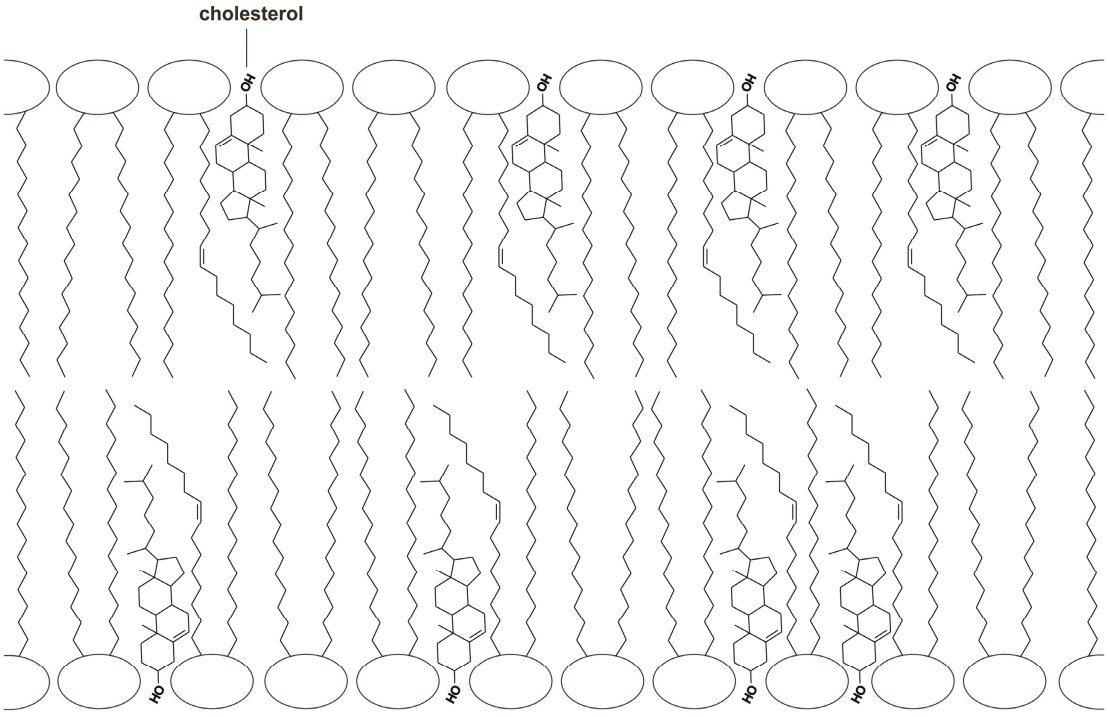


Fig. 2. Overview of alignment of cholesterol in the lipid bilayer of membranes of eukaryotic cells. Adapted from Wolfe (1993) and Alberts *et al*. (1994).

(Wolfe, 1993; Alberts *et al*., 1994; Fig. 2). Furthermore, the presence of sterols helps preventing crystallization of hydrocarbons and phase shifts in cell membranes.
The membranes of most bacteria, both Gram-positive and Gram-negative, do not contain sterols. Some bacteria, however, do contain hopanoids in their membranes (Ourisson *et al*., 1979). Hopanoids are structurally related to sterols; both hopanoids and sterols are derived from the precursor squalene (Bode *et al*., 2003). In some bacteria, the presence of sterols was reported, e.g. in *Rhodopseudomonas palustris* and some methylotrophic bacteria and cyanobacteria (Aaronson, 1964; Bode *et al*., 2003; Volkman, 2003). These, however, are exceptions to the rule, and there is ongoing debate whether the sterols found indeed originate from the bacteria themselves or whether they were artifacts from sample preparation.

A large variety of sterol molecules, differing in their side-chain or steroid ring system, is found in Nature. For example, sterol side chains differ in length, may possess C24 branches, and/or contain unsaturated bonds or epoxy substituents. The steroid ring system may harbor modifications of the 3-hydroxy moiety (e.g. glycosylation, acetylation or oxidation to 3-keto), vary in the degree of saturation or contain additional hydroxyl or keto functional groups at various positions.
The largest variety of sterol molecules is found in plants; more than 250 different phytosterols have been identified (Akihisa *et al*., 1991). The most abundant phytosterols are β-sitosterol (Fig. 1), campesterol and stigmasterol. Plant sterols may constitute up to 5% of the total lipid content in seeds and vegetable oil (Engel and Knorr, 2004). The main sterol found in fungi is ergosterol, but additional sterols have been identified in e.g. yeast (zymosterol and cerevisterol) (Honeywell and Bills, 1932). In higher animals only cholesterol is present while in some lower animals, including sponges, also modified forms of cholesterol have been identified. In addition to its important structural role in membranes, cholesterol also serves as precursor for the synthesis of all steroids, bile acids and vitamin D in animals (Berg *et al*., 2002).

MICROBIAL STEROID RING DEGRADATION

As early as 1913, Söhngen reported on the ability of a *Mycobacterium* strain to utilize phytosterols as sole carbon and energy source. Since then, many more bacterial species and strains have been described able to degrade steroids. Especially *Actinobacteria* are known for their capability to degrade sterols and steroids, of which members of the genera *Mycobacterium*, *Rhodococcus*, *Nocardia* and *Arthrobacter* stand out as being efficient sterol degraders. Table 1 shows an extensive list of *Actinobacteria* that are capable of sterol side chain transformation.
Also some Gram-negative bacteria are able to metabolize steroids or bile acids, including *Pseudomonas* spp. (Tenneson *et al*., 1979; Dhar and Samantha, 1993) and *Comamonas testosteroni* (formerly known as *Pseudomonas testosteroni*) (Levy and Talalay 1959).

The oxidative pathways involved in steroid and sterol degradation have been studied at the biochemical and genetic level using the aforementioned bacteria. A global overview of microbial cholesterol degradation is depicted in Fig. 3. Transcriptomic analysis of *R. jostii* RHA1 has identified a large set of genes that were upregulated during growth of strain RHA1 on cholesterol: a gene cluster of 51 upregulated genes, comprised within *ro04482-ro04706* (Fig. 4), codes for degradation of the greater part of cholesterol (van der Geize *et al*., 2007, [Chapter 2]). Based on these data, and the combined results described above, we proposed a catabolic pathway for these compounds (van der Geize *et al*., 2007, [Chapter 2]). During the microbial degradation of sterols, steroids like 4-androstene-3,17-dione (AD) or 1,4-androstadiene-3,17-dione (ADD) may be formed as pathway intermediates (Fig. 3) that are further degraded by dedicated enzymes (van der Geize *et al*., 2007, [Chapter 2]). The enzymatic reactions involved in steroid nucleus degradation are of pharmaceutical interest, as they enable modification of steroids in a regio- and stereo selective manner, which is much more difficult to achieve chemically.

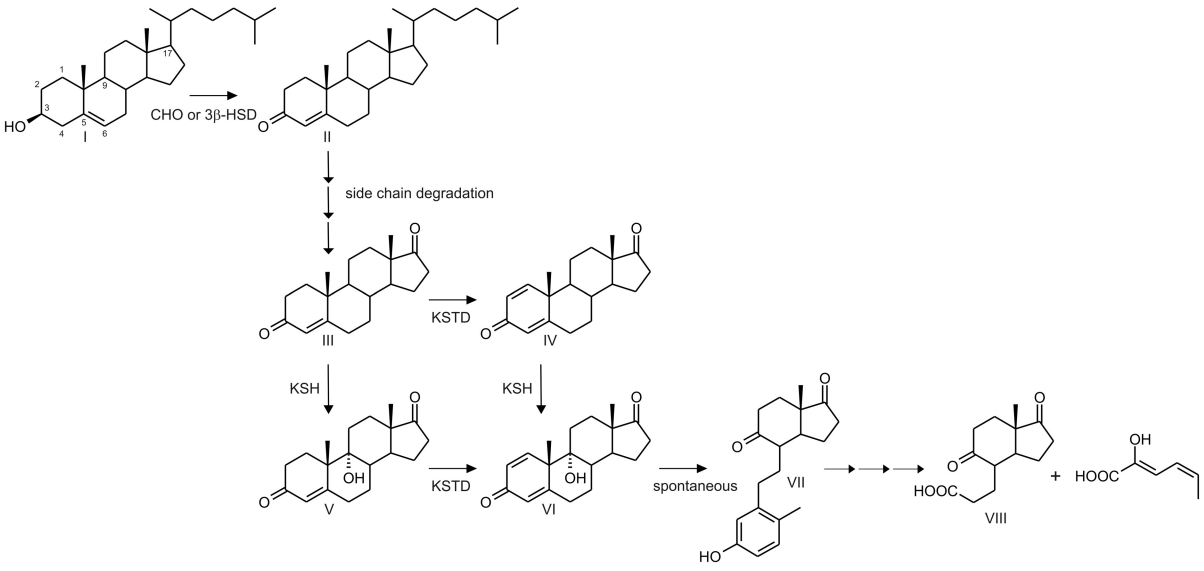


Fig. 3. The deduced pathway of microbial cholesterol (compound I) nucleus opening. Steroid nucleus degradation is initiated by oxidation of the 3-hydroxyl group into 3-keto, followed by isomerization of the C5-6 (Δ5) double bond to C4-5 (Δ4), yielding 4-cholestene-3-one (compound II), catalyzed by either cholesterol oxidase (CHO) or 3β-hydroxysteroid dehydrogenase (3β-HSD). Next (or prior to CHO/3β-HSD activity, depending on the bacterial strain/species), the C17-alkyl side chain is removed via β-oxidation ultimately leading to formation of 17-keto steroids, like 4-androstene-3,17-dione (AD, compound III). Two key steps in the degradation of the AD steroid nucleus are Δ1-dehydrogenation, mediated by 3-ketosteroid Δ1-dehydrogenase (KSTD), and 9α-hydroxylation, catalyzed by 3-ketosteroid 9α-hydroxylase (KSH) with formation of either 1,4-androstadiene-3,17-dione (ADD, compound IV) or 9-hydroxy-4-androstene-3,17-dione (9OH-AD, compound V) as intermediates. The ultimate product of the combined reactions (9OH-ADD) is unstable (compound VI) and spontaneously rearranges into 3-hydroxy-9,10-secoandrosta-1,3,5,(10)-triene-9,17-dione (3-HSA, compound VII). Further oxidation leads to formation of 9,17-dioxo-1,2,3,4,10,19-hexanorandrostan-5-oic acid (DOHNAA, compound VIII) and eventually CO_2 and H_2O.

The initial steps of 3-ketosteroid ring degradation are Δ1-dehydrogenation and 9α-hydroxylation, involving KSTD (3-ketosteroid Δ1-dehydrogenase) and KSH (3-ketosteroid 9α-hydroxylase) enzymes (Fig. 3). The order of these reactions is unclear and may be different among bacterial species or strains. The product ultimately formed by the joint action of these reactions (9OH-ADD) is chemically instable and leads to spontaneous opening of ring B and aromatization of ring A. In Fig. 3 the pathway of steroid ring degradation is depicted. In a further sequence of reactions, rhodococci and *M. tuberculosis* completely degrade the steroid molecule (to CO_2 and H_2O). In macrophages, *M. tuberculosis* has also been shown to accumulate propionyl-CoA from cholesterol, which is converted into branched chain lipids (Pandey and Sassetti, 2008, Yang *et al*., 2009).

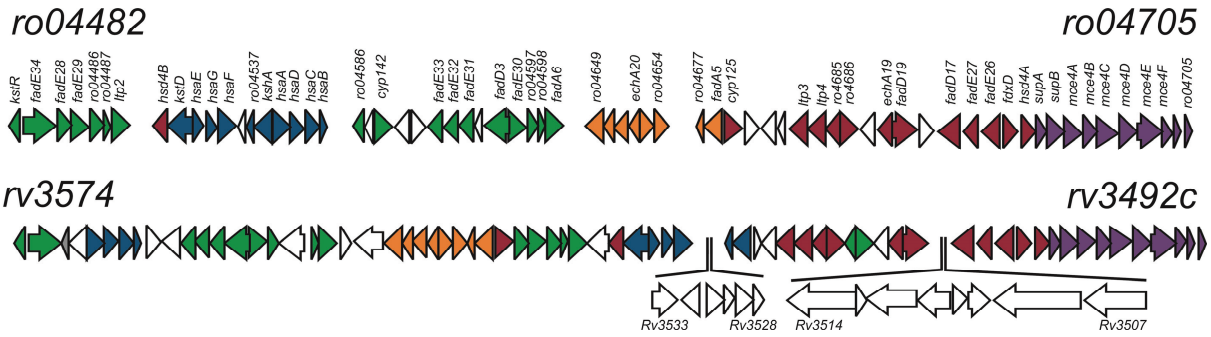


Fig. 4. The cholesterol catabolic gene cluster of *R. jostii* RHA1 comprising *ro04482* (= *kstR*) - *ro046706* (above) and *M. tuberculosis* H37Rv (below). The genes are color-coded according to their (predicted) function: purple, uptake; blue, cleavage of rings A and B; orange, degradation of 9,17-dioxo-1,2,3,4,10,19-hexanorandrostan-5-oic acid (DOHNAA); green, degradation of rings C and D (van der Geize *et al*., 2007).

3-Ketosteroid Δ1-dehydrogenases

3-Ketosteroid Δ1-dehydrogenases (EC 1.3.99.4; KSTD) are flavoenzymes that catalyze the deprotonation of C1(α) and C2(β) from the A ring of 3-ketosteroid nuclei (Itagaki *et al*., 1990a). The electrons are transported to the flavin moiety, but the final acceptor remains elusive. The first KSTD activity was described in the testosterone degrader *C. testosteroni*, using intact cells (Levy and Talalay, 1959a) and crude extracts (Levy and Talalay, 1959b), while the first purified and characterized KSTD enzyme was from *Nocardia restrictus* (Sih and Bennet, 1962). The first cloned and sequenced *kstD* gene originated from *C. testosteroni* (Plesiat *et al*., 1991). More recently, *kstD* homologs were identified and characterized in *R. erythropolis* SQ1 (van der Geize *et al*., 2000, 2001; Knol *et al*., 2008) and *Mycobacterium smegmatis* (Brzostek *et al*., 2005). Strain SQ1 has at least three *kstD* genes (van der Geize *et al*., 2000, 2001; Knol *et al*., 2008), while the genome of *R. jostii* RHA1 encodes five putative *kstD*s (McLeod *et al*., 2006; van der Geize *et al*., 2007, [Chapter 2]; Knol *et al*., 2008). It is still unclear why so many KSTD isoenzymes are present in these rhodococci. Interestingly, the genome of *M. tuberculosis* contains only one *kstD* and also single copies of the *kshA* and *kshB* genes (see below). Possibly, the relatively constant conditions inside their host cells

(e.g. pH and temperature) and the limited variety of steroid substrates available may explain why this obligatory pathogen contains only a single isoform of these steroid ring degradation activities (KstD and KshA). Conversely, the diverse and varying environments inhabited by free living rhodococci may require a larger array of genes encoding steroid degradation activities to ensure maximal efficiency of steroid degradation under a wide variety of conditions. However, more in depth biochemical studies on KSTD (and KSH) enzymes are required to verify this.

3-Ketosteroid 9α-hydroxylases

3-Ketosteroid 9α-hydroxylases (EC 1.14.99.24; KSH) are two-component enzymes belonging to the class IA monooxygenases that catalyze hydroxylation of the C9 atom of steroids. Together with KSTD, KSH activity is required for steroid ring opening as depicted in Fig. 3. KSH enzymes of *Rhodococcus* consist of an iron-sulphur containing hydroxylase, KshA, and an flavin containing reductase part, KshB (van der Geize *et al*., 2002b; Petrusma *et al*., 2009). Studies on KSH activity have long been hampered by difficulties to obtain purified, active proteins, mainly due to oxygen lability (Chang and Sih, 1964; Strijewski, 1982; van der Geize *et al*., 2002). Only very recently, the first report on heterologously expressed and purified KSH from *R. rhodochrous* DSM43269 has been published (Petrusma *et al*., 2009). Interestingly, strain DSM43269 was shown to encode five *kshA* genes and only deletion of all five *kshA*s resulted in complete blockage of growth on AD (Petrusma *et al*., in preparation; Wilbrink *et al*., submitted, [Chapter 4]). Further research is needed to elucidate the specific roles of the different KshA homologs in this strain. The genome of *M. tuberculosis* H37Rv contains only one *kshA* gene. $KshA_{H37Rv}$ was purified and its crystal structure has been elucidated (Capyk *et al*., 2009). Furthermore $KshA_{H37Rv}$ is important for survival of *M. tuberculosis* in macrophages (Rengarajan *et al*., 2005) and was shown to be a pathogenicity factor (Hu *et al*., 2009).

Further steroid ring oxidation

The combined action of KSTD and KSH leads to opening of steroid ring B and aromatization of ring A, forming 3-hydroxy-9,10-secoandrosta-1,3,5(10)-triene-9,17-dione (3-HSA) from AD (Fig. 3). 3-HSA is then hydroxylated at the C4 position to a catechol (Horinouchi *et al*., 2004). This reaction is catalyzed by a two component monooxygenase system consisting of TesA1A2 in *C. testosteroni* (Horinouchi *et al*., 2004) and HsaAB in *R. jostii* RHA1 and *M. tuberculosis* (van der Geize *et al*., 2007, [Chapter 2]; Dresen *et al*., 2010). The catechol is then further degraded via a meta-cleavage pathway, involving HsaC in *R. jostii* RHA1 and *M. tuberculosis*, to yield 3,4-dihydroxy-9,10-secoandrosta-1,3,5(10)-triene-9, 17-dione (3,4-DHSA) (van der Geize *et al*., 2007, [Chapter 2]; Yam *et al*., 2009) and subsequent C-C bond hydrolysis takes place, mediated by HsaD (Lack *et al*., 2009, 2010). The same reactions are catalyzed by TesB and TesD, respectively, in *C. testosteroni* (Horinouchi *et al*., 2001; 2003; 2004). The further degradation of the remaining steroid ring nucleus has not been studied at the genetic or biochemical level. In strain RHA1, *ro04483-ro04488* and *ro04586-ro04599*,

located within the cholesterol catabolic gene cluster, were suggested to be responsible for degradation of rings C and D (van der Geize *et al*., 2007, [Chapter 2]).

MICROBIAL STEROL DEGRADATION

Sterol uptake: the mce4 cluster

The cholesterol catabolic gene cluster of strain RHA1 comprises a sterol uptake system encoded by the *mce4* cluster (*ro04696-ro04705*), whose transcription was upregulated during growth of strain RHA1 on cholesterol as compared to pyruvate grown cells (van der Geize *et al*., 2007, [Chapter 2]). The cluster consists of an operon of 11 genes that encodes two permeases (*supA* and *supB*) and the Mce4A-Mce4I proteins that penetrate the outer layer of the mycolic acids and possibly are involved in substrate binding, together constituting a complex ATP-binding cassette (ABC) transporter system. The ATPase domain, however, was neither encoded by the *mce4* locus, nor was it found elsewhere in the strain RHA1 cholesterol catabolic gene cluster (van der Geize *et al*., 2007, [Chapter 2]; Mohn *et al*., 2008). The *mceG* gene of *M. tuberculosis* encodes an ATPase domain belonging to the Mkl family which has been shown to interact with proteins encoded by the *mce1* and *mce4* loci, but is not located proximal to either of the two loci (Joshi *et al*., 2006). In strain RHA1, *ro01974* and *ro02744* encode MceG orthologs, therefore, either one of them or both may be involved in the Mce4 transport system (Mohn *et al*., 2008). The Mce4 transport system mediates sterol uptake specifically, since mutagenesis studies with genes from the *mce* cluster revealed that they are essential for growth of strain RHA1 on sterols, including cholesterol, but not other steroids (Mohn *et al*., 2008). It cannot be ruled out that the Mce4 system also transports various steroids, but most likely other mechanism(s) for uptake of steroids exist that are either specific for steroids or that can complement steroid uptake in strain RHA1 *mce4* gene deletion mutants (Mohn *et al*., 2008).

Cholesterol oxidase and 3β-hydroxysteroid dehydrogenase

An important early reaction in microbial sterol degradation is the formation of a 3-keto-4-ene ring structure, catalyzed by cholesterol oxidase or 3β-hydroxysteroid dehydrogenase. These enzymes are both able to oxidize the β-hydroxyl group at C3 with subsequent isomerization of the Δ5 double bond to a Δ4 (Fig. 3).

Cholesterol oxidase (CHO, E.C. 1.1.3.6) enzymes are flavoproteins that can either contain a covalently or non-covalently bound FAD (Croteau and Vrielink, 1996; Wilmańska *et al*., 1995; Vrielink and Ghisla, 2009). Furthermore, they may function extracellularly or inside the cell, depending on the type of enzyme and organism (Wilmańska *et al*., 1995). Cholesterol oxidases transfer the hydrogen atoms from various steroid or sterol substrates directly to molecular oxygen, thereby forming hydrogen peroxide. Also, some cholesterol oxidases are known to perform hydroxylation of cholesterol, ultimately forming 4-cholesten-6β-ol-3-one from cholesterol (Molnár *et al*., 1993; Biellmann, 2001; Doukyu *et al*., 2009). Many *Actinobacteria* produce cholesterol oxidase enzymes, including members of *Rhodococcus*

(Aihara *et al*., 1986; Navas *et al*., 2001; Fernández de Las Heras *et al*., 2010), *Brevibacterium* (Fujishiro *et al*., 1990) and *Streptomyces* (Murooka *et al*., 1986; Ishizaki *et al*., 1989).
Also 3β-hydroxysteroid dehydrogenases (3β-HSD, E.C. 1.1.1.145) are known to catalyze 3-keto-4-ene formation (Fig. 3), in *Nocardia* sp. (Horinouchi *et al*., 1991) and *M. tuberculosis* (Yang *et al*., 2007). These enzymes use $NAD(P)^+$ as electron acceptor instead of molecular oxygen and function intracellularly. The 3β-HSD enzyme Rv1106c of *M. tuberculosis* was shown to be active on cholesterol, pregnenolone and dehydroepiandrosterone, while the highest activities were found for the latter two compounds (Yang *et al*., 2007). Pregnenolone and dehydroepiandrosterone are C21 and C19 steroids, respectively, and are expected pathway intermediates of cholesterol degradation. Therefore, it is likely that in *M. tuberculosis* cholesterol side chain degradation occurs prior to ring oxidation. Recent mutational and biochemical studies have shown that in *R. jostii* RHA1 C26 hydroxylation is the obligate first step in cholesterol degradation, prior to the action of CHO or 3β-HSD, while in *R. rhodochrous* DSM43269 there was no clear preference for either of these two reactions (Rosłoniec *et al*., 2009, [Chapter 3]).

Microbial sterol side chain degradation
The mechanism of microbial degradation of the 17-alkyl side chains of sterols has been elucidated several decades ago, by identifying reaction intermediates that accumulated in natural or UV-mutagenized strains of *Nocardia* and *Mycobacterium* that were blocked in various steroid nucleus degradation steps. Furthermore, crude cell extracts of natural or mutant strains, capable of sterol degradation, were used to assess several enzyme steps of sterol side chain shortening at the biochemical level. In conclusion, it was deduced that sterol side chain degradation proceeds via a process similar to β-oxidation (Sih *et al*., 1968a, 1968b; Arima *et al*., 1969; Marsheck *et al*., 1972; Fujimoto *et al*., 1982a, 1982b). The general mechanism of β-oxidation and the enzymatic steps involved are shown below (Fig. 5). Although the mechanism of sterol side chain degradation is well understood, very limited information is available on the genes and purified enzymes involved in the process: virtually all experiments were done using crude extracts from strains whose genomic sequence data is unavailable. The subject of microbial sterol side chain degradation is relevant for applications in the pharmaceutical industry and has been reviewed regularly (Martin, 1977; Kieslich, 1985; Szentirmai, 1990; Fernandes *et al*., 2003).

Formation of sterol C26 carboxylic acid
It has been shown that in *Mycobacterium* sp. 2104 cholesterol side chain degradation is initiated by C26 hydroxylation and is preceded by 3-keto-4-ene formation of the nucleus (Zaretskaya *et al*., 1968). We recently demonstrated that the rhodococcal P450 monooxygenase CYP125 is essential for C26 carboxylic acid formation and most likely functions as a C26-hydroxylase (Rosłoniec *et al*., 2009, [Chapter 3]). Subsequent studies on the CYP125 homolog of *M. tuberculosis* H37Rv have confirmed that $CYP125_{H37Rv}$ catalyzes C26 or C27 hydroxylation (Capyk *et al*., 2009; McLean *et al*., 2009) and further oxidation to

yield a C26/27-oic acid (Ouellet *et al*., 2010; Johnston *et al*., 2010). The high amino acid sequence identities between the CYP125 enzymes from strain H37Rv and strains RHA1 and DSM43269 strongly suggests that also the rhodococcal enzymes catalyze the full oxidation of sterol C26 to C26-oic acid. Subsequent degradation of the side chain of the sterol C26-oic acid occurs via β-oxidation (Fig. 5).

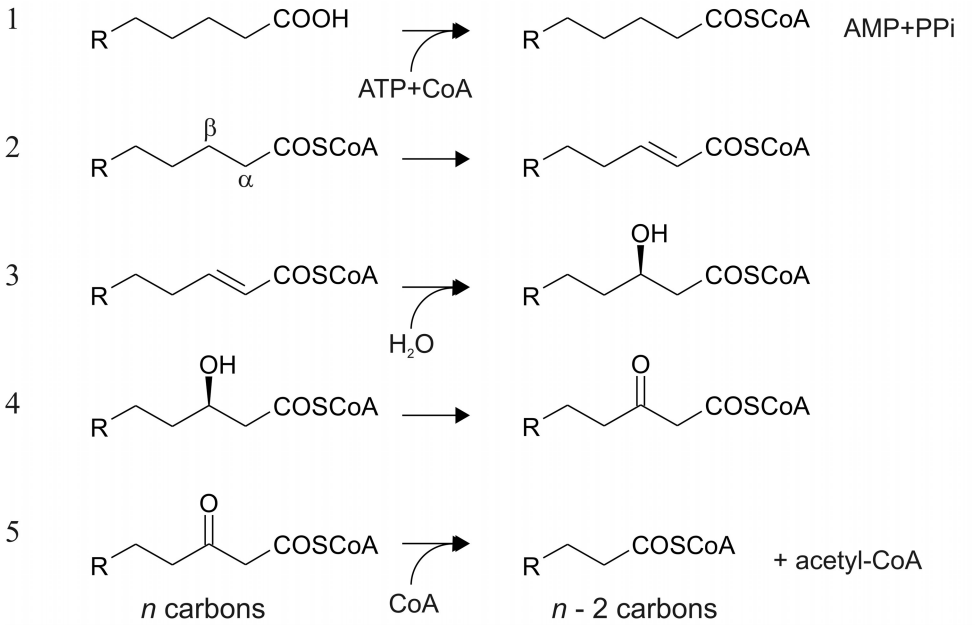


Fig. 5. Schematic overview of the enzymatic steps involved in β-oxidation: 1) CoA activation of the terminal carboxylic acid at the expense of ATP; a two-step reaction involving adenylation and thioesterification, ultimately leading to an acyl-CoA product, catalyzed by CoA ligase, 2) dehydrogenation of the α and β carbons, catalyzed by acyl-CoA dehydrogenase, 3) hydration of the double bond between α and β, catalyzed by enoyl-CoA hydratase, 4) second dehydrogenation step, yielding a 3-ketoacyl-CoA intermediate, catalyzed by β-hydroxyacyl-CoA dehydrogenase, 5) thiolytic cleavage; cleavage of 3-ketoacyl-CoA by a second CoA molecule yielding acetyl-CoA and an acyl-CoA shortened by two carbon atoms, catalyzed by thiolase. Depending on the chain-length of the substrate several consecutive rounds of β-oxidation may occur, catalyzed by the same enzyme set.

Sterol side chain activation

The side chains of sterols like cholesterol are degraded in a stepwise manner, by a process similar to β-oxidation of fatty acids (Fig. 6A). The first reaction of β-oxidation is activation of the sterol carboxylic acid moiety with CoA, catalyzed by steroid-CoA ligase. CoA ligase catalyzed reactions are driven by ATP hydrolysis and require Mg^{2+} as a cofactor. CoA activation of acyl substrates is a two-step process involving an enzyme-bound adenylated intermediate (Groot *et al*., 1976) and thioester formation, where the AMP is replaced by CoA (Chang *et al*., 1997).

A 65 kDa CoA ligase from *Mycobacterium* sp. NRRL B-3805 was purified to near homogeneity and shown to be highly specific towards C26-carboxylic acid steroids (Chen, 1985). The CoA ligase was shown to be present as a single enzyme, while the rest of the β-

oxidation enzymes were aggregated, forming a loosely bound complex (Chen, 1985). The gene encoding the steroid-CoA ligase, or any of the other β-oxidation enzymes involved, remained unidentified.

The *baiB* gene from *Eubacterium* sp. VPI 12708 was shown to encode a CoA ligase involved in cholic acid degradation (Mallonee *et al*., 1992). FadD19 of *R. rhodochrous* DSM43269, displaying only 24% amino acid sequence identity to the *baiB* gene product, was shown to act as a CoA ligase involved in sterol side chain degradation (Wilbrink *et al*., submitted, [Chapter 4]). Using *R. rhodochrous* strain mutant RG32, devoid of steroid 9α-hydroxylation activity but capable of sterol side chain degradation, it was shown that deletion of *fadD19*$_{DSM43269}$ resulted in complete abolishment of β-sitosterol side chain degradation, but not of cholesterol (Wilbrink *et al*., submitted, [Chapter 4]). FadD19$_{DSM43269}$ was able to activate the side chain of cholesterol 26-oic acid (both 3-hydroxy-5-ene and 3-oxo-4-ene forms) *in vitro*, suggesting that the enzyme may also be involved in cholesterol side chain

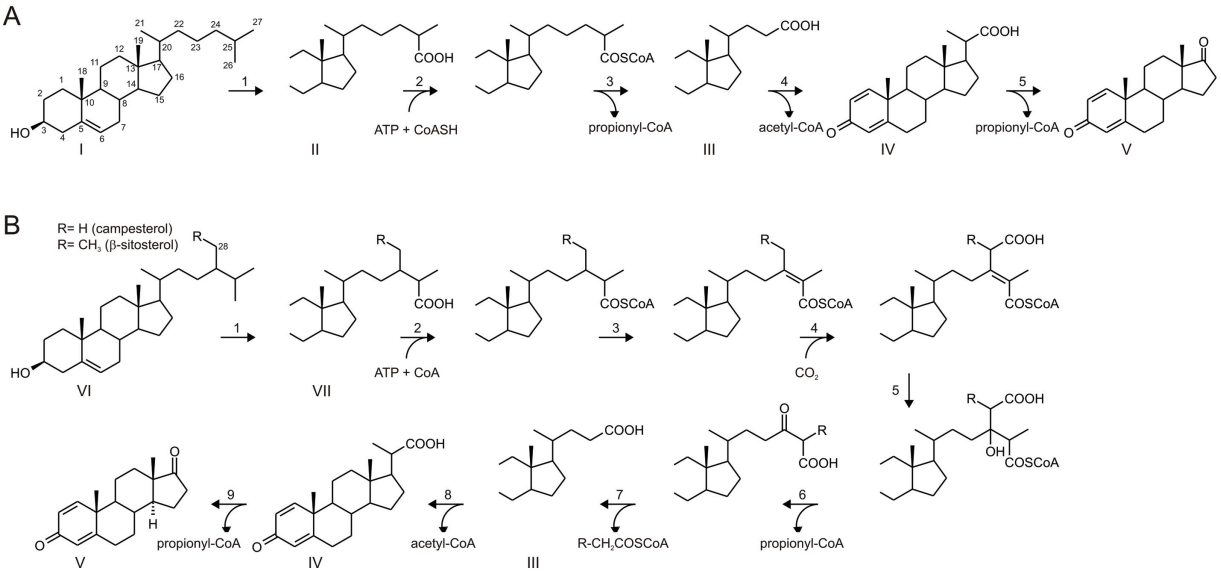


Fig. 6. Overview of the degradation pathways of the side chain of cholesterol (A) and the C24-branched side chains of β-sitosterol and campesterol (B) in *Actinobacteria* (adapted from Sih *et al*., 1968b and Fujimoto *et al*., 1982b, respectively). Cholesterol side chain degradation is initiated by (1) formation of a cholesterol C26-oic acid, involving CYP125 (Rosłoniec *et al*. 2009, [Chapter 3]). Next, the carboxylic acid is activated by a CoA ligase (2) (Wilbrink *et al*., submitted, [Chapter 4]) and propionyl-CoA is released via a β-oxidation enzyme complex (3). Acetyl-CoA is subsequently released after a second round of β-oxidation (4), followed by aldolytic cleavage of C19-20 (5).
The side chain of β-sitosterol is also oxidized at position C26 (1), followed by CoA activation by CoA ligase FadD19 (2) (Wilbrink *et al*., submitted, [Chapter 4]). A dehydrogenase introduces a double bond between C24-25 (4), which is followed by carboxylation of C28 (5). After hydration of the double bond (6), propionyl-CoA is released through an aldol lyase reaction (7) most likely involving *ltp3* and/or *ltp4* (Chapter 5). A CoA activation step (8) is then followed by two rounds of β-oxidation (9) and formation of a C19 steroid by aldolytic cleavage (10), identical to cholesterol side chain degradation as shown above.

degradation (Wilbrink *et al*., submitted, [Chapter 4]). However, since deletion of *fadD19*$_{DSM43269}$ in mutant RG32 did not affect cholesterol side chain degradation, (an)other CoA ligase(s) most likely is/are encoded by the strain DSM43269 genome; such CoA ligase(s) involved in cholesterol degradation remain to be identified.

β-oxidation enzyme complex

Further degradation of the CoA-activated sterol side chain has been demonstrated in crude extracts, involving a multi-enzyme complex consisting of an acyl-CoA dehydrogenase or oxidase, enoyl-CoA hydratase, 3β-hydroxyacyl-CoA dehydrogenase and thiolase (Chen, 1985). In *Mycobacterium* sp. NRRL B-3805, the first dehydrogenation step is performed by an FMN dependent acyl-CoA oxidase (Chen, 1985).

In *Pseudomonas* sp. Chol1, an acyl-CoA dehydrogenase (ACAD) has been identified that catalyzes dehydrogenation of cholyl-CoA (Birkenmaier *et al*., 2007). BLAST search of the ACAD of strain Chol1 against the *R. jostii* RHA1 genome shows that strain RHA1 contains several genes encoding close homologs of this protein, including *ro05825* and *fadE26* (= *ro04693*), both with 50% amino acid sequence identity to ACAD. These genes are located within clusters that are upregulated during growth of strain RHA1 on cholic acid (unpublished) and cholesterol, respectively (van der Geize *et al*., 2007, [Chapter 2]). Deletion of the *fadE26* homolog in *R. rhodochrous* RG32, however, did not affect cholesterol side chain degradation (Wilbrink *et al*., submitted, [Chapter 4]). Furthermore, the cumulative inactivation of various genes encoding acyl-CoA dehydrogenases (*fadE26*, *fadE27* and *ro04690*$_{DSM43269}$) did not abolish cholesterol side chain degradation in strain RG32 (Wilbrink *et al*., submitted, [Chapter 4]). This suggests that multiple isoenzymes are capable to catalyze such dehydrogenation reactions.

The next step in side chain degradation is hydration of the double bond by an enoyl-CoA hydratase. However, no enzymes or genes encoding this activity are presently known. The strain RHA1 cholesterol catabolic gene cluster contains *echA19*, a gene encoding a putative enoyl-CoA hydratase which, due to its location proximal to other genes involved in side chain degradation, was suggested to be responsible for the double bond hydration of sterol side chains (van der Geize *et al*., 2007, [Chapter 2]). However, this remains to be confirmed by experimental studies.

After the hydration step, another dehydrogenation reaction is performed, catalyzed by 3β-hydroxyacyl-CoA dehydrogenase. In *Mycobacterium* sp. NRRL-B3805, this reaction was shown to require NAD^+ as a cofactor, but the enzyme has not been characterized in more detail (Chen, 1985). The sterol side chain degradation locus of strain RHA1 contains *hsd4A*, a putative 3β-hydroxyacyl-CoA dehydrogenase that was highly upregulated during growth on cholesterol (van der Geize *et al*., 2007, [Chapter 2]). This gene is the most likely candidate to encode the second dehydrogenation step in cholesterol side chain degradation. However,

no molecular or biochemical characterization studies have been performed on Hsa4A yet. Therefore, the role of *hsa4A* in cholesterol metabolism remains to be elucidated.

The final reaction of a β-oxidation cycle is thiolytic cleavage of the substrate. In the case of cholesterol side chain degradation, thiolase mediated C-C bond cleavage of carbons 24-25 and 22-23 results in formation of C24 and C22 steroids and release of propionyl-CoA and acetyl-CoA, respectively (Sih *et al*., 1968a, 1968b; Fujimoto *et al*., 1982a). Removal of the remaining three carbon side chain of the C22 steroid intermediate is then catalyzed by a mechanism different from β-oxidation, most likely involving a reverse aldol-lyase reaction (Fujimoto *et al*., 1982a). An alternative pathway has been proposed for *Mycobacterium aurum*, involving a decarboxylation step, thus forming progesterone from 3-oxo-23,24-bisnor-4-cholene-22-oic acid (4-BNC) (Prome *et al*., 1983). Further degradation may then proceed by an oxidation reaction catalyzed by a Baeyer-Villiger monooxygenase (BVMO) and hydrolysis by an esterase, respectively. In *R. rhodochrous*, a steroid BVMO has been identified that catalyzes oxidation of progesterone into testosterone acetate (Miyamoto *et al*., 1995; Morii *et al*., 1999). In strain RHA1, the steroid BVMO encoding gene homolog (*ro02492*), however, was shown not to be upregulated during growth on cholesterol (van der Geize *et al*., [Chapter 2]). Therefore, it remains unclear how the side chain of 4-BNC or 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC) is removed in strain RHA1.

Mutational analysis revealed that *fadA5* from *M. tuberculosis* encodes a thiolase that is essential for cholesterol side chain degradation to form AD and ADD (Nesbitt *et al*., 2009). The ability of FadA5 to degrade acetoacetyl-CoA *in vitro* confirmed that the enzyme is indeed a thiolase (Nesbitt *et al*., 2009). In *R. rhodochrous* RG32, however, inactivation of the *fadA5* homolog did not affect cholesterol degradation, despite the very high similarities between their amino acid sequences. Also deletion of two other genes, *ltp3* and *ltp4* in *R. rhodochrous* strain RG32, annotated as nonspecific lipid carrier protein and thiolase, respectively, did not affect cholesterol degradation (Chapter 5). It appears likely that the genome of *R. rhodochrous* encodes many more thiolases, similar to observations made with sequenced *Rhodococcus* genomes, and redundancy or enzymatic compensation may explain why cholesterol degradation still occurs when *fadA5* has been inactivated.

C24-branched chain sterol side chain degradation

Degradation of C24-branched chain sterols, like the phytosterols β-sitosterol, campesterol and the fungal ergosterol, requires several additional enzymatic steps compared to cholesterol. Fujimoto *et al*. (1982b) used a cell-free system derived from *Mycobacterium* sp. NRRL B-3805 to demonstrate that the C24-branched side chains of β-sitosterol and campesterol are carboxylated at the C28 position, following C26 oxidation and CoA activation (Fig. 6B). The carboxylase responsible for incorporation of HCO_3^- at the C28 position was not inhibited by avidin, an effective inhibitor of biotin dependent enzymes. This carboxylase thus appears to lack biotin or this prosthetic group is deeply buried inside

the enzyme (Chen, 1985). Upon formation of the carboxylate intermediate, the action of enoyl-CoA hydratase is followed by cleavage of C24-C25 by an aldol-lyase, resulting in release of propionyl-CoA from β-sitosterol and campesterol (Fig. 6B, Chen, 1985; Fujimoto *et al*., 1982b). Interestingly, deletion of *ltp3* or *ltp4* in *R. rhodochrous* strain RG32 resulted in complete blockage of β-sitosterol side chain degradation, but not of that of cholesterol (Chapter 5). Both *ltp3* and *ltp4* encode proteins with similarities to the thiolase SCPx. However, closer bioinformatic analysis revealed that two thiolase signature motifs, comprising the highly conserved catalytic residues, were absent in Ltp3 and Ltp4. Furthermore, due to the structure of C24-branched chain sterols, formation of a ketoacyl-CoA, required for thiolase activity, is chemically impossible and degradation of these compounds was suggested to occur via aldolytic cleavage similar to degradation of other tertiary β-hydroxy CoA esters. (Fujimoto *et al*., 1982b) Therefore, *ltp3* and *ltp4* were suggested to encode aldol lyase type enzymes with a role in C-C cleavage of C24-branched chain sterols (Chapter 5). This is the first report on the identification and characterization of enzymes (Ltp3 and Ltp4) with specific and essential roles in carbon-carbon cleavage of branched chain sterols in *Rhodococcus* strains.

Transcriptional regulation of steroid degradation genes

Transcription of cholesterol metabolic genes in *Actinobacteria* is tightly regulated. The first regulatory gene involved in steroid degradation has been identified in *Arthrobacter simplex* and named *ksdR* (Molnár *et al*., 1995). A second gene, identified in *R. erythropolis* SQ1 and denoted *kstR*, showed little similarity with *ksdR* at the amino acid level (van der Geize *et al*., 2000). In both cases, these regulatory genes were located upstream of *kstD* genes. In *M. smegmatis*, a TetR-type transcriptional repressor encoded by *kstR1* was shown to control the expression of a large regulon involved in sterol metabolism (Kendall *et al*., 2007). Deletion of *kstR1* in *M. smegmatis* resulted in upregulation of 132 genes, (6- to 1771-fold), while 27 genes were downregulated (6- to 18-fold). The *kstR1* gene is highly conserved among mycobacteria and other closely related *Actinobacteria* like *N. farcinica* and *R. jostii* RHA1 (Kendall *et al*., 2007; van der Geize *et al*., 2007, [Chapter 2]). Recently, a second TetR-type repressor, named *kstR2*, was identified in *M. tuberculosis* and shown to control expression of a regulon consisting of 15 genes with a suggested role in cholesterol metabolism (Kendall *et al*., 2010). Also *kstR2* is highly conserved among sterol degrading *Actinobacteria*.

Cholesterol degradation and pathogenicity

Part of the success of *M. tuberculosis* as a pathogen lies in its remarkable ability to invade human macrophages, and to reside and even to replicate in this hostile environment (Flynn and Chan, 2001). Together, these properties make it very hard to combat infections by *M. tuberculosis* strains. Several studies have identified a large array of genes that are essential for survival of *M. tuberculosis* H37Rv in macrophages, using amongst others genome wide transposon site hybridization (TraSH) studies (Sassetti and Rubin, 2003; Sassetti *et al*., 2003;

Rengarajan *et al*., 2005; Joshi *et al*., 2006). Of the genes identified, some were unique to mycobacteria. The physiological roles of many of these genes, however, remained unknown (Schnappinger *et al*., 2003; Rengarajan *et al*., 2005). The cholesterol catabolic gene cluster identified in the closely related actinomycete *R. jostii* RHA1 was found to be highly conserved in strain H37Rv (van der Geize *et al*., 2007, [Chapter 2]). Many of the strain H37Rv homologs encoded by the cluster were previously found to be required for macrophage survival and could now be identified as genes involved in cholesterol degradation (van der Geize *et al*., 2007, [Chapter 2]). Cholesterol had previously been known to be essential for phagocytosis of *M. tuberculosis* by macrophages, and for inhibition of macrophage maturation, but the mechanism(s) involved remained elusive in these studies (Gatfield and Pieters, 2000; Peyron *et al*., 2000; de Chastellier and Thilo, 2006).

Recently, it was confirmed that *M. tuberculosis* is indeed able to grow on cholesterol as carbon and energy source, and that this metabolic feature is important for persistence in macrophages (Pandey and Sassetti, 2008; Brzosteck *et al*., 2009). Recent studies also have confirmed that many of the enzymes involved in cholesterol degradation are essential for survival of *M. tuberculosis* in macrophages; hence these are pathogenicity factors, including KSH (Hu *et al*., 2009), CYP125 (Chang *et al*., 2009), cholesterol oxidase (Brzostek *et al*., 2007), KSTD (Rengarajan, *et al*., 2005), HsaD (Lack *et al*., 2009) and FadA5 (Nesbitt *et al*., 2010). TraSH studies previously showed that H37Rv homologs of the rhodococcal sterol CoA ligase FadD19 (Wilbrink *et al*., submitted [Chapter 4]) and Ltp3-Ltp4 (Chapter 5) were not essential for survival in macrophages (Rengarajan *et al*., 2005). Since the enzymes involved in actinobacterial cholesterol metabolism generally have no human counterparts, they are suitable targets for the development of inhibitors which may ultimately lead to new anti-TB drugs.

AIM OF THE RESEARCH

The aim of my PhD research was to identify and characterize genes and enzymes involved in actinobacterial sterol side chain degradation. Elucidation of the exact mechanism, and knowledge of the genes/enzymes involved in this process, may ultimately lead to the cell engineering of bacterial strains with enhanced steroid production from sterols, or construction of mutants that accumulate sterol intermediates with partially degraded side chains which may serve as building blocks for novel steroids. The research focused mainly on members of the bacterial genus *Rhodococcus*, efficiently degrading sterols, further stimulated by experience with an extensive genetic toolbox in our laboratory, as well as the availability of genome sequences of various rhodococci.

The research project was funded by the NWO/ACTS program Integration of Biosynthesis and Organic Synthesis (IBOS).

SCOPE OF THIS THESIS

Chapter 1 provides a general introduction to this thesis, describing properties of *Actinobacteria*, with an emphasis on the genus *Rhodococcus*. The structures and sources of steroids and sterols, and the genes and enzymes involved in their degradation by *Actinobacteria*, are discussed. The biotechnological interest in the synthesis of bioactive steroids from relatively cheap natural sterols is briefly reviewed, as well as the possible treatment of *Mycobacterium tuberculosis* infections by interfering with its cholesterol catabolism, essential for pathogenesis and survival in macrophages.

Chapter 2 describes the identification of a cholesterol catabolic gene cluster in *R. jostii* RHA1 by transcriptome analysis. The genes from this cluster were annotated and physiological roles for the encoded proteins in degradation of the sterol molecule were proposed, including those predicted to be involved in side chain degradation. Biochemical experiments were performed with some of the enzymes encoded by the cluster, to confirm their role in sterol degradation. Also mutational analyses and qRT-PCR were done to assess the role of specific genes in sterol catabolism. A cluster of 51 genes (*ro04482-ro04706*, Fig. 4), encoding degradation of rings A, B and likely C and the 17-alkyl side chain of cholesterol, was found to occur in a 235 kb stretch of DNA in the strain RHA1 genome. Bioinformatic analysis revealed that homologs of the genes are conserved and organized in one contiguous cluster in *M. tuberculosis* H37Rv. Strikingly, many of the homologs from strain H37Rv were previously shown to be important in survival of this pathogen in macrophages, but their function hitherto had remained unknown. Therefore, the results presented here provide insight into the role of cholesterol metabolism in *M. tuberculosis* infection, and potential targets for novel biotherapeutics.

Chapter 3 describes the identification of CYP125 as a sterol C26-hydroxylase, essential in the formation of sterol C26-carboxylic acid intermediates. Previous results in *R. jostii* RHA1 showed that during growth on cholesterol, *cyp125* (= *ro04679*) was one of the most highly upregulated genes from the cholesterol catabolic gene cluster. This chapter shows that a deletion mutant of *cyp125* in strain RHA1 was blocked in growth on 3-hydroxysterols, but

not on 3-oxo derivatives. Purified $CYP125_{RHA1}$ was shown to bind sterols, but no enzyme activity could be demonstrated. Inactivation of *cyp125* in *R. rhodochrous* strain RG32 resulted in the complete blockage of sterol side chain degradation activity, strongly suggesting that this gene encodes the sterol 26-hydroxylase. Gene deletion mutant RHA1Δ*cyp125* indeed was able to grow on 5-cholestene-26 oic acid-3β-ol, the C26-carboxylic acid derived from cholesterol and the predicted product of CYP125. Gene disruption mutant RG32Ω*cyp125* was able to convert the side chain of 5-cholestene-26 oic acid-3β-ol, but not cholesterol, showing that CYP125 is essential for the conversion of cholesterol into the C26-oic acid catabolite during sterol side chain degradation.

Chapter 4 reports the molecular characterization of FadD19 of *R. rhodochrous* DSM43269 as a steroid CoA ligase with an essential and specific role in degradation of C24-branched chain sterols. Deletion of *fadD19* in strain RG32 resulted in complete abolishment of degradation of the C24-branched side chains of β-sitosterol and campesterol, whereas degradation of the side chain of cholesterol remained unimpaired. Heterologously produced $FadD19_{DSM43269}$ was shown to be active towards 5-cholestene-26 oic acid-3β-ol and 3-oxo-4-cholestene-26 oic acid, confirming that the enzyme is a steroid-CoA ligase.

Chapter 5 describes the essential roles of *ltp3* and *ltp4* in degradation of C24-branched chain sterols. The *ltp3* and *ltp4* (*ro04683* and *ro04684* in strain RHA1, respectively) gene orthologs from *R. rhodochrous* DSM43269 were cloned. Deletion of either *ltp3* or *ltp4* in *R. rhodochrous* strain RG32 resulted in a blockage of β-sitosterol side chain degradation, but not of cholesterol. Functional complementation showed that the observed phenotypes were caused by the absence of the respective genes, rather than by polar effects. Both *ltp3* and *ltp4* encode proteins with sequence similarity to thiolases. However, since removal of C24-branched sterol side chains does not involve thiolytic cleavage, their gene products most likely have aldol lyase activity.

Chapter 6 provides a summary and discussion of the results presented in this thesis, and an outlook for future research.

REFERENCES

References are listed on pages 119-140.

CHAPTER 2

A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages

Robert van der Geize, Katherine Yam, Thomas Heuser, Maarten H. Wilbrink, Hirofumi Hara, Matthew C. Anderton, Edith Sim, Lubbert Dijkhuizen, Julian E. Davies, William W. Mohn and Lindsay D. Eltis

Published in Proc. Natl. Acad. Sci. USA (2007) 104: 1947-1952

ABSTRACT

Rhodococcus sp. strain RHA1, a soil bacterium related to *Mycobacterium tuberculosis*, degrades an exceptionally broad range of organic compounds. Transcriptomic analysis of cholesterol-grown RHA1 revealed a catabolic pathway predicted to proceed via 4-androstene-3,17-dione (AD) and 3,4-dihydroxy-9,10-seconandrost-1,3,5(10)-triene-9,17-dione (3,4-DHSA). Inactivation of each of the *hsaC*, *supAB* and *mce4* genes in RHA1 substantiated their roles in cholesterol catabolism. Moreover, the *hsaC*$^{-}$ mutant accumulated 3,4-DHSA, indicating that $HsaC_{RHA1}$, formerly annotated as a biphenyl-degrading dioxygenase, catalyzes the oxygenolytic cleavage of steroid ring A. Bioinformatic analyses revealed that 51 rhodococcal genes specifically expressed during growth on cholesterol, including all predicted to specify the catabolism of rings A and B, are conserved within an 82-gene cluster in *M. tuberculosis* H37Rv and *Mycobacterium bovis* bacillus Calmette-Guérin. *M. bovis* bacillus Calmette-Guérin grew on cholesterol, and *hsaC* and *kshA* were up-regulated under these conditions. Heterologously produced $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$ transformed 3,4-DHSA and its ring-cleaved product, respectively, with apparent specificities ~40-fold higher than for the corresponding biphenyl metabolites. Overall, we annotated 28 RHA1 genes and proposed physiological roles for a similar number of mycobacterial genes. During survival of *M. tuberculosis* in the macrophage, these genes are specifically expressed, and many appear to be essential. This is the first time a complete suite of genes necessary for microbial steroid degradation has been delineated and the first time pathogenic mycobacteria have been shown to catabolize cholesterol. The results suggest that cholesterol metabolism is central to *M. tuberculosis*'s unusual ability to survive in macrophages and provide insights into potential targets for novel therapeutics.

INTRODUCTION

Rhodococci are a genus of GC-rich, mycolic acid-producing bacteria within the order *Actinomycetales* that includes *Mycobacterium* (Gurtler *et al*., 2004). Rhodococci degrade a broad range of organic compounds, particularly hydrophobic ones, thereby playing a key role in the global carbon cycle. Analysis of the 9.7 Mb genome of RHA1 (www.rhodococcus.ca) reveals that this organism harbors a diverse armamentarium of enzymes (McLeod *et al*., 2006), consistent with the catabolic versatility of the genus. These catabolic activities, together with robust and rapid rhodococcal growth, are of great interest to pharmaceutical, environmental, chemical and energy industries (van der Geize *et al*. 2004).

The bacterial catabolism of steroids has attracted considerable attention (van der Geize *et al*. 2004) in part as a potential means of producing bioactive steroids from natural, low-cost sterols such as β-sitosterol and cholesterol. A pathway responsible for the aerobic degradation of the latter via 4-androstene-3,17-dione (AD) and 3-hydroxy-9,10-seconandrost-1,3,5(10)-triene-9,17-dione (3-HSA) may be pieced together from biochemical and genetic studies in diverse bacteria (Fig.1). In some *Mycobacterium* (Wovcha *et al*., 1978) and *Rhodococcus* (Murohisa and Ida, 1993a, 1993b) species, the aliphatic side chain at C17 is removed via a process similar to β-oxidation involving progressively shorter carboxylic acids. In these strains, 3-ketosteroid Δ1-dehydrogenase (KSTD) and 3-ketosteroid 9α-hydroxylase (KshAB) catalyze the opening of ring B and aromatization of ring A to yield 3-HSA (van der Geize *et al*. 2001, 2002b, 2004; Gibson *et al*., 1966). The subsequent degradation of 3-HSA to 9,17-dioxo-1,2,3,4,10,19-hexanorandrostan-5-oic acid (DOHNAA) via oxygenolytic cleavage of ring A is specified by the *tes* genes in the testosterone-degrading strain *Comamonas testosteroni* TA441 (Horinouchi *et al*., 2003; 2004). In *Rhodococcus equi*, the propionate moiety of DOHNAA is removed via β-oxidation (Miclo and Germain, 1992). Many of the genes involved in steroid catabolism have yet to be identified, and many of the pathway enzymes are poorly characterized, particularly those involved in degrading the bicycloalkanone originating from rings C and D. Detailed knowledge of steroid catabolism is essential to engineering strains for the biotransformation of sterols.

Recent genomic analyses revealed that rhodococci may be useful models for many mycobacterial processes: approximately 60% of the 3999 genes of *Mycobacterium tuberculosis* H37Rv are conserved in RHA1, including many of unknown function (McLeod *et al*., 2006). *M. tuberculosis* is the leading cause of mortality from bacterial infection, killing 2-3 million people worldwide each year, and extensive drug resistant strains such as XDR-TB are now emerging (WHO, 2005). One poorly characterized aspect of mycobacterial physiology that contributes to the prevalence of tuberculosis (TB) is the bacterium's unusual ability to survive for long periods of time, and even to replicate, in the normally hostile environment of the macrophage (Zhang, 2005; Clark-Curtiss and Haydel, 2003). The mechanisms enabling this persistence are poorly understood, but are logical targets for novel therapeutic agents. Transposon site hybridization (TraSH), a genome-wide microarray-based technique, identified 126 genes that appear to be necessary for survival of H37Rv in

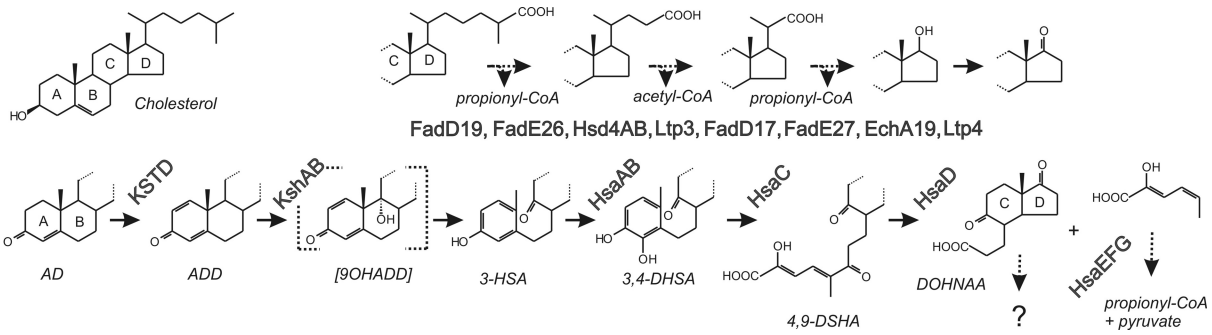


Fig. 1. The deduced cholesterol catabolic pathway of *Rhodococcus* sp. RHA1, *M. tuberculosis* H37Rv and *M. bovis* bacillus Calmette-Guérin. The enzymatic steps of side chain degradation and ring opening are depicted. The latter are important for H37Rv survival in the macrophage (Fig. 2). Dashed arrows indicate multiple enzymatic steps. The compound in brackets undergoes non-enzymatic hydrolysis. Genes responsible for the degradation of rings C and D in RHA1 are not conserved in H37Rv nor bacillus Calmette-Guérin. ADD, 1,4-androstadiene-3,17-dione; 9OHADD, 9α-hydroxy-1,4-androstadiene-3,17-dione; KshAB, 3-ketosteroid 9α-hydroxylase.

macrophages under conditions that model the immune response (Rengarajan *et al*., 2005) and many others that are critical for *in vivo* survival in mice (Sassetti and Rubin, 2003). Further, transcriptomic studies have identified suites of genes that are specifically up-regulated during survival in the macrophage (Schnappinger *et al*., 2003). Despite the importance of these genes, their physiological roles are largely unknown.

We investigated the cholesterol catabolic pathway in RHA1 by comparing the transcriptomes of cholesterol- and pyruvate-grown cells. Targeted gene deletion was used to substantiate key catabolic steps. Bioinformatic analyses enabled annotation of many of the cholesterol catabolic genes and also revealed their presence in *M. tuberculosis* and *M. bovis*. Conditions to grow *M. bovis* on cholesterol were developed and the expression of 2 pathway genes was shown by qRT-PCR. Two of the *M. tuberculosis* pathway enzymes were heterologously produced and shown to efficiently catalyze the predicted transformations of steroid ring A. The results are discussed with respect to the survival of *M. tuberculosis* in the macrophage.

RESULTS

The cholesterol transcriptome of RHA1

In liquid medium containing 2 mM cholesterol as sole organic substrate, RHA1 grew to a density of 2 x 10^8 cells/ml. Microarray analysis revealed 572 genes that were up-regulated at least 2-fold during growth on cholesterol compared to on pyruvate. Many of the up-regulated genes are scattered throughout the 9.7-Mb genome (www.rhodococcus.ca) and likely reflect a general physiological adaptation of the bacterium to growth on a highly hydrophobic, polycyclic compound. However, six clusters of up-regulated genes were clearly discerned (supplemental Table S1). The most striking of these was a cluster of 51 genes that occur within a 235 Kb-stretch of RHA1's 7.9-Mb chromosome (*ro04482-ro04705*; Fig.2A). As discussed below, these genes encode proteins with significant sequence identity with

enzymes involved in the catabolism of steroid rings A and B by *C. testosteroni* TA441 (Horinouchi *et al*., 2003, 2004) and *Rhodococcus erythropolis* SQ1 (van der Geize *et al*. 2001, 2002b). A second cluster of chromosomal genes (*ro06687-ro06698*) also appear to be involved in cholesterol catabolism. The four other gene clusters (*ro00440-ro00453*, *ro03461-ro03464*, *ro08053-ro08060* and *ro10126-ro10162*) do not appear to be directly involved in steroid catabolism, and some are described elsewhere (Gonçalves *et al*., 2006).

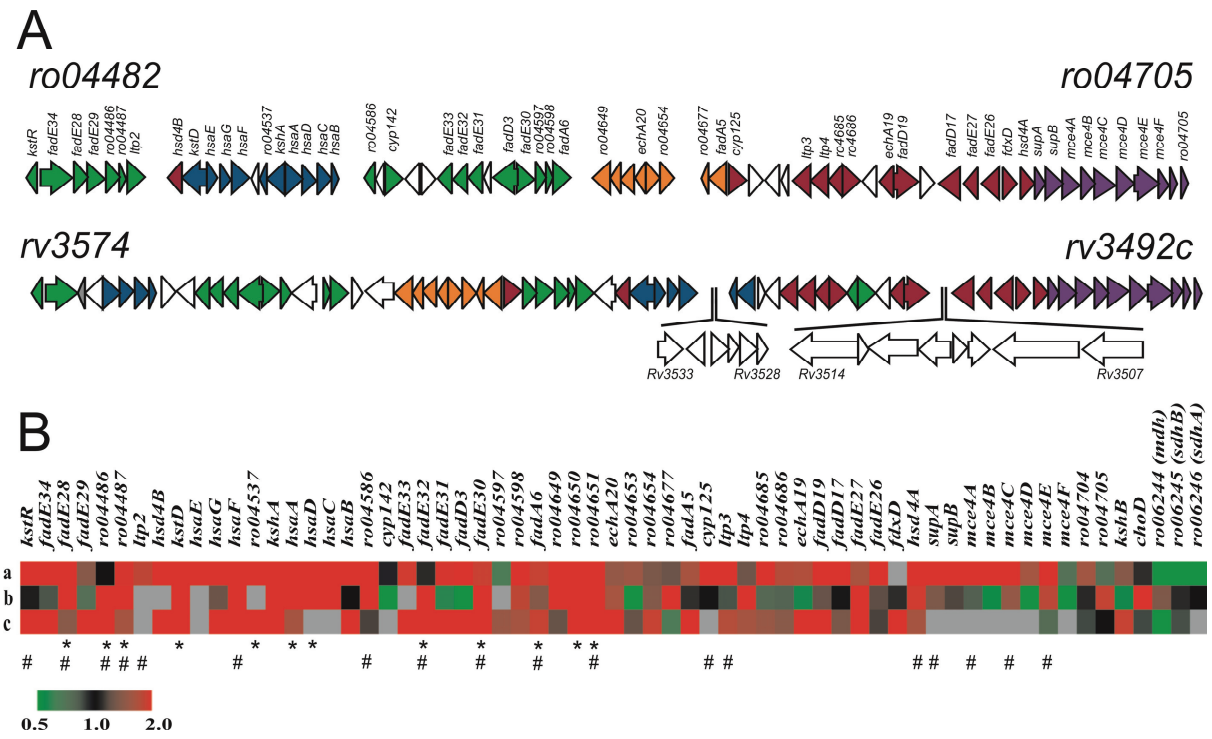


Fig. 2. The cholesterol catabolic genes of *Rhodococcus sp*. RHA1 and *M. tuberculosis* H37Rv: comparison of their organization and of their activities in different studies. (A) Genes in the physical map are color-coded according to assigned function: purple, uptake; red, side-chain degradation; blue, cleavage of rings A and B; orange, degradation of the DOHNAA propionate moiety; green, degradation of rings C and D. White arrows represent genes for which no reciprocal homologue is present. The nucleotide sequences of the *M. tuberculosis* H37Rv and *M. bovis* bacillus Calmette-
-Guérin clusters share 96% identity. (B) Heat map indicating correlation between (a) gene expression (fold difference) during growth of RHA1 on cholesterol versus pyruvate, (b) effect of gene disruption on H37Rv survival in IFN-γ-activated macrophages according to TraSH analysis (reciprocal of ratio) (Rengarajan *et al*., 2005), (c) gene expression in H37Rv after 48 h of growth in IFN-γ-activated macrophages (Schnappinger *et al*., 2003). *M. tuberculosis* genes predicted as essential for survival in the macrophage (Pethe *et al*., 2004; Rengarajan *et al*., 2005) and *in vivo* in mice (Sassetti and Rubin, 2003) are indicated with * and #, respectively.

Annotation of cholesterol catabolic genes

Among the genes that were up-regulated during growth on cholesterol, the annotation of those predicted to specify cholesterol catabolism is summarized in supplemental Table S2. Most of these comprise the 51 genes of the *ro04482-ro04705* cluster (Fig. 2A) and most of

the encoded proteins have sufficient sequence similarity to well characterized enzymes that their function can be confidently predicted. Thus, sequences of KshA, KshB and KSTD (Fig. 1) are 40-69% identical to those of orthologs in *R. erythropolis* SQ1 (supplemental Table S2) that act sequentially to transform AD to 3-HSA (van der Geize *et al*. 2001, 2002b, 2004). Further degradation of 3-HSA was predicted to be specified by 7 genes, annotated here as *hsa*, that are clustered with *kstD* and *kshA* (Figs. 1, 2A). The encoded proteins share significant amino acid sequence similarity (30-60%, supplemental Table S2) with the *tes*-encoded enzymes of *C. testosteroni* TA441 that transform 3-HSA during growth on testosterone (Horinouchi *et al*., 2003, 2004). HsaC and HsaD were previously annotated as BphC5 and BphD2 in RHA1, respectively, due to the former's ability to catalyze the extradiol cleavage of 2,3-dihydroxybiphenyl (DHB), and their sequence similarity to the corresponding biphenyl catabolic enzymes (Sakai *et al*., 2002). However, HsaC shares greater sequence identity with TesB of *C. testosteroni* TA441 (Horinouchi *et al*., 2004) than with extradiol dioxygenases that preferentially cleave DHB. Moreover, qRT-PCR analyses confirmed that *hsaC* was up-regulated 15.4-fold during growth of RHA1 on cholesterol as compared to on either biphenyl or pyruvate.

It was more difficult to assign specific roles to the numerous β-oxidation genes of the *ro04482-ro04705* cluster. Most of these gene products share greatest sequence identity with homologs that occur in *M. tuberculosis* H37Rv and were annotated accordingly (supplemental Table S2). One set of these genes (*hsd4A*, *hsd4B*, *fadD19*, *fadE26* and *ltp3*) is highly up-regulated and encodes all the enzymes necessary to perform one full cycle of β-oxidation. Hsd4A and Hsd4B share intriguing sequence similarity with the eukaryotic multifunctional 17β-hydroxysteroid dehydrogenase IV (17βHSD4) involved in peroxisome related disorders (Mindnich *et al*., 2004). Hsd4A is homologous to the N-terminal domain of 17βHSD4, which acts as a 17β-hydroxysteroid dehydrogenase and, with branched fatty acids and bile acids, as a D-3-hydroxyacyl-CoA dehydrogenase. Hsd4B is homologous to the central domain of 17β-HSD4, which is a 2-enoyl acyl-CoA hydratase proposed to be involved in cholesterol side chain shortening. Accordingly, we predict that these RHA1 genes specify at least one cycle of β-oxidative transformation of the C17 side chain to propionyl-CoA and acetyl-CoA. A second near complete set of β-oxidation genes (*echA19*, *fadD17*, *fadE27* and *ltp4*) are up-regulated to a lesser extent, but are likely also involved in side-chain degradation. The bi-functional Hsd4A likely transforms the 17β-hydroxysteroid resulting from cleavage of the cholesterol side chain.

A third cluster of up-regulated genes related to β-oxidation, including *fadE28*, is similar to those involved in testosterone catabolism by *C. testosteroni* TA441 (Horinouchi *et al*., 2003; 2004). These genes may be involved in the degradation of the DOHNAA originating from steroid rings C and D (Fig. 1), as this part of the molecule is common to testosterone and cholesterol whereas the C17 side chain is not. This set of genes is preceded by a gene encoding a TetR-type transcriptional regulator similar to the KstR (32% identity) repressor of *kstD* (van der Geize *et al*. 2001), suggesting that the RHA1 genes are also regulated by steroids. The propionate moiety of DOHNAA is likely degraded by β-oxidation encoded by

the gene cluster that includes *echA20*. This cluster includes genes encoding a two-subunit, ATP-dependent CoA transferase of the type thought to initiate β-oxidation (Miclo and Germain, 1990).

The *ro04482-ro04705* cluster also includes eight genes that appear to encode a multi-component cholesterol uptake system: *supAB* and *mce4ABCDEF*. Conserved Domain data revealed the presence of a domain related to an ABC-transport system involved in resistance to organic solvents in both SupA and SupB (Marchler-Bauer *et al*., 2003). The *mce* cluster is one of two such clusters in RHA1 that are highly similar to the four sets of "mammalian cell entry" (*mce*) genes of *M. tuberculosis* H37Rv (Cole *et al*., 1998). Mce proteins are critical virulence factors in *M. tuberculosis* (Rengarajan *et al*., 2005) although the exact role of these genes is unknown. Heterologously expressed *mce1A* enhanced the entry of *E. coli* into non-phagocytic HeLa cells (Arrud *et al*., 1993), while *mce1*$^{-}$ and *mce4*$^{-}$ strains of *M. tuberculosis* H37Rv showed attenuated survival in mice (Joshi *et al*., 2006). It has been proposed that Mce proteins are components of transport systems that translocate lipids between the bacterial cell and its host (Joshi *et al*., 2006). Consistent with this, Mce1A is expressed at the cell surface of *M. tuberculosis* H37Rv (Chitale *et al*., 2001). Indeed, signal sequences are predicted for all of the Mce4 proteins of RHA1 except Mce4C (SignalP (Bendtsen *et al*., 2004)), indicating that these are secreted or surface-exposed proteins. In summary, the 51 up-regulated genes of the *ro04482-ro04705* cluster appear to include all of those necessary to specify the catabolism of cholesterol to DOHNAA.

The separate *ro06687-ro06698* gene cluster, induced on cholesterol, includes genes typical of those encoding cycloalkanone catabolism (supplemental Table S2). These include *ro06698* and *ro06693,* which encode a probable monooxygenase and lactone hydrolase, respectively. We predict that these genes are involved in degrading steroid ring D of DOHNAA.

Annotation of the cholesterol catabolic genes further revealed that these genes are but one of four sets in RHA1 that appear to specify the catabolism steroid-like compounds. Each of these sets encodes homologs of all ring-degrading enzymes: KshAB, KstD, HsaAB, HsaC, HsaD, and at least one cyclohexanone monooxygenase. Sequence analyses revealed that all the KshA homologs (ro02490, ro04538, ro05811 and ro09003) share at least 52% amino acid sequence identity with KshA of *R. erythropolis* SQ1 (van der Geize *et al*. 2002b). Phylogenetic analyses (supplemental Fig. S1A) revealed that these enzymes define a subclass of Rieske non-heme oxygenases. Similarly, all the HsaC homologs (ro02488, ro04541, ro05803 and ro09005) share at least 37% amino acid sequence identity and key active site residues with TesB of *C. testosteroni* TA441. These enzymes constitute a previously unrecognized subclass of type I extradiol dioxygenases (supplemental Fig. S1B) distinct from those involved in biphenyl and naphthalene catabolism. Similar analyses of HsaA and HsaD revealed comparable relationships (data not shown): for each type of enzyme, the known steroid-degrading homologs constitute a distinct subclass. None of the additional three sets of genes were up-regulated in RHA1 during growth on cholesterol and so appear to encode degradation of other steroids.

Mutational analysis of cholesterol catabolic genes

The critical role of Mce4A-Mce4F and SupAB proteins in cholesterol catabolism was confirmed by unmarked in-frame gene deletion of the entire *mce4ABCDEF* gene cluster and the *supAB* genes, respectively, in RHA1. Both the *mce4* and *sup* mutants were severely impaired in the ability to grow on cholesterol in liquid mineral medium (Table 1). By contrast, growth on AD was not affected, supporting our hypothesis that Mce4 and SupAB are specifically involved in the uptake of cholesterol in RHA1. The doubling times of RHA1 and the mutants on AD (~12 hours) was approximately three times longer than on pyruvate or benzoate.

To substantiate the predicted role of HsaC in catalyzing the extradiol cleavage of 3,4-DHSA, a catechol, *hsaC* was deleted. In liquid media, the *hsaC*$^-$ mutant grew on cholesterol at a rate that was 60% that of the wild-type strain and developed a pink color. By contrast, growth on pyruvate was not affected. The slower growth on cholesterol may be due either to degradation of the C17 side chain or complementary activity of one of the HsaC homologs in RHA1 (supplemental Fig. S1B). The pink color is consistent with the accumulation and non-enzymatic oxidation of a catechol. To identify the latter, metabolites were extracted from the supernatant of *hsaC*$^-$ cells incubated in the presence of cholesterol. HPLC analysis revealed a major metabolite which, when derivatized with TMS, yielded a compound with a molecular ion m/z = 460 (supplemental Fig. S2). The molecular ion and its fragmentation pattern correspond to those predicted for TMS-derivatized 3,4-DHSA. Finally, transformation of the metabolite with $HsaC_{H37Rv}$ as described below yielded a product with a pH-dependent spectrum essentially identical to that reported for 4,9-DSHA (Gibson *et al*., 1966) (ε_{392} = 7.64 $mM^{-1}cm^{-1}$ at pH 8.0), confirming the metabolite's identity as 3,4-DHSA (Fig. 1).

Table 1. Growth yields of RHA1 and mutants on different organic substrates.

	Cholesterol (1 mM)	AD (1 mM)	Pyruvate (20 mM)	Benzoate (20 mM)
WT	73 (5)	77 (12)	170 (20)	470 (60)
Δ*supAB*	1 (1)	75 (11)	170 (20)	520 (50)
Δ*mce4*	3 (1)	78 (14)	150 (40)	440 (90)

Growth yields are expressed as µg protein per ml of culture medium and are averages of triplicate cultures. Values in parentheses are standard errors.

Conservation of the cholesterol catabolic pathway in mycobacteria

Further bioinformatic analyses revealed that 58 genes of the *ro04482-ro04705* cluster in RHA1, including the 51 that were up-regulated during growth on cholesterol, are conserved together with much of their putative operonic structure within an 82-gene cluster in the genomes of *M. tuberculosis* H37Rv (*Rv3492c-Rv3574*; Fig. 2A) and *M. bovis* bacillus

Calmette-Guérin (*Bcg3556c-Bcg3639*; www.sanger.ac.uk/Projects/M_bovis/) as well as within an 80-gene cluster in *M. avium* (*subsp*. *paratuberculosis*) (*Map0571-Map0491*; Li *et al*., 2005). As noted above, these genes appear to be sufficient to specify the uptake of cholesterol, the β-oxidation of the branched side chain at C17, and the catabolism of rings A and B to central metabolites via 3-HSA to yield DOHNAA (Fig. 1). The sequence identities of the RHA1, H37Rv and bacillus Calmette-Guérin homologs are summarized in supplemental Table S2. Phylogenetic analyses revealed that among the 4 sets of steroid-degrading enzymes in RHA1, the mycobacterial enzymes are most similar to those involved in cholesterol catabolism (supplemental Fig. S1).

Cholesterol catabolism in *M. bovis* bacillus Calmette-Guérin

Initial attempts to grow bacillus Calmette-Guérin on cholesterol as sole energy source met with limited success, as for other pathogenic mycobacteria (Av-Gay and Sobouti, 2000). However, using a liquid minimal medium containing asparagine, citrate, and Triton (Schnappinger *et al*., 2003), the final growth yield of bacillus Calmette-Guérin was proportional to the initial concentration of cholesterol in the medium. Thus, in medium supplemented with 0, 0.25 mM and 0.5 mM cholesterol, respectively, the overall protein yields were 22 ± 7, 46 ± 9, and 70 ± 4 μg/ml. Further modification of the medium to reflect host factors or to improve the availability of the cholesterol to the bacterium may improve growth.

To investigate whether the predicted cholesterol catabolic pathway is involved in this growth of bacillus Calmette-Guérin, qRT-PCR analyses were performed on *kshA* and *hsaC* using *sigA* as a control. Normalized transcript levels were significantly higher in cultures growing on cholesterol (n = 4) than on glucose (n = 5) for both of *kshA* ($p < 0.005$) and *hsaC* ($p < 0.05$), with relative fold-differences of 3.7 and 2.4, respectively. Similar results were observed when comparing cholesterol- to pyruvate-grown cells. The relative fold differences for *kshA* and *hsaC* are very similar to the expression ratios determined for these genes (4.6 and 2.1, respectively) using the microarray to compare RHA1 growing on cholesterol versus on pyruvate (supplemental Table S1); although, a slightly higher fold difference was determined for *hsaC* in RHA1 using qRT-PCR. The relative fold differences for *kshA* and *hsaC* are also very similar to those determined for these genes using a microarray to compare H37Rv growing in macrophages versus *in vitro* (Schnappinger *et al*., 2003).

The catalytic activities of $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$

To substantiate the predicted cholesterol catabolic pathway in *M. tuberculosis* H37Rv, the activities of two central enzymes were investigated, $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$. These enzymes were targeted in part because they were previously annotated as putative biphenyl-degrading enzymes (Sakai *et al*., 2002). Accordingly, $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$ were heterologously expressed in *E. coli* and their steady-state kinetic parameters were evaluated using cell extracts. As summarized in Table 2, the enzymes preferentially transformed cholesterol metabolites as compared to biphenyl metabolites. Specifically, cell extracts

Table 2. Steady-state kinetic parameters of $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$ for steroid and biphenyl metabolites.

Enzyme	Substrate	K_m (μM)	V_{max} (μM s^{-1})	V_{max}/K_m (s^{-1})
$HsaC_{H37Rv}$	3,4-DHSA	0.9 (0.5)	12 (4)	790 (370)
	DHB	8.5 (0.8)	2.5 (0.4)	18 (3)
$HsaD_{H37Rv}$	4,9-DSHA	4 (1)	0.06 (0.02)	1.0 (0.2)
	HOPDA	19 (6)	0.009 (0.003)	0.028 (0.007)

Parameters were normalized to the amount of cellular extract (mg of protein content) used in the assays. Values in parentheses represent standard errors.

containing $HsaC_{H37Rv}$ catalyzed the extradiol cleavage of 3,4-DHSA with an apparent specificity 44-fold higher than for DHB. Similarly, extracts containing $HsaD_{H37Rv}$ catalyzed the hydrolysis of 4,9-DSHA with an apparent specificity 34-fold higher than for HOPDA. Equivalent extracts prepared from cells that contained the empty vector did not detectably transform either the steroid or biphenyl metabolites. These results strongly support the predicted roles of the mycobacterial enzymes in steroid metabolism and also indicate that the aliphatic side chain of cholesterol is removed prior to ring degradation.

DISCUSSION

The current study identified clusters of genes that encode the catabolism of cholesterol in RHA1. These were initially identified through bioinformatic analyses of genes that were up-regulated during growth on cholesterol. Moreover, the involvement of Mce4 and SupAB proteins in cholesterol catabolism and the role of $HsaC_{RHA1}$, an extradiol dioxygenase, were substantiated using gene deletion and characterization of the resultant mutants. Steroids such as cholesterol are ubiquitous in plants, animals and some microbes and likely comprise an important energy source for saprophytic bacteria, particularly actinomycetes which efficiently use hydrophobic substrates. While various aspects of steroid catabolism have been described in different bacteria (Gibson *et al*., 1966; Horinouchi *et al*., 2003, 2004; Miclo and Germain 1990, 1992; Murohisa and Ida, 1993a, 1993b; van der Geize *et al*. 2001, 2002b; Wovcha *et al*., 1978), the current study is the first in which the genes of an entire catabolic pathway are delineated in a single organism. The identified genes include several involved in sterol uptake and side chain degradation that were not previously identified and that are particularly good targets for cell and enzyme engineering studies. Thus, sterol uptake is believed to be rate-limiting yet its mechanism is poorly understood. Similarly, efficient sterol side chain degradation is critical for high yield production of steroid intermediates, particularly as most sterols used in microbial transformations consists of a mixtures of compounds with slightly different side chains that are transformed with

different efficiencies. Overall, this study facilitates the development of whole cell biotransformation processes for the synthesis of industrially relevant steroid compounds.

A second important contribution of the current study is the discovery that the cholesterol catabolic pathway is conserved in related pathogenic actinomycetes including *M. tuberculosis*, *M. bovis* and *M. avium*. Thus, the latter appear to have retained the capacity for cholesterol metabolism and exploited it to survive in their hosts. Consistent with our bioinformatic predictions, *M. bovi*s bacillus Calmette-Guérin used cholesterol as a carbon and energy source and genes encoding the ring-degrading enzymes KshA and HsaC were up-regulated during this utilization. The substrate of the pathway in *M. tuberculosis* was verified by demonstrating the apparent specificity of $HsaC_{H37Rv}$, an extradiol dioxygenase, and $HsaD_{H37Rv}$, a C-C bond hydrolase, for the steroid metabolites. These enzymes had been annotated as a probable DHB dioxygenase (NP_218085) and a HOPDA hydrolase (CAB07143), respectively. The current study further suggests that these enzymes do not play a direct role in mycobacterial cell wall synthesis as recently suggested (Anderton *et al*., 2006). Of the pathway proteins conserved in RHA1 and *M. tuberculosis*, those with the lowest amino acid sequence identities are the Mce4 proteins. It is possible that the latter have different functions in the two organisms. However, our findings in RHA1 are consistent with the recent proposal that the *supAB* and *mce4* genes encode a lipid-transport system (Joshi *et al*., 2006). Moreover, some of the genes that were functionally linked to this system in that study include several cholesterol catabolic genes. Thus, the low sequence identities of the RHA1 and mycobacterial Mce4 homologs may instead reflect the different environments from which these two strains must scavenge cholesterol.

Several lines of evidence indicate that the identified steroid catabolic pathway is essential for the survival of *M. tuberculosis* in the macrophage. First, 41 of the pathway genes, including those specifying catabolism of rings A and B, are among those specifically up-regulated during survival in the macrophage (Fig. 2B, (Schnappinger *et al*., 2003)). Second, TraSH analyses predict that at least 11 of the pathway genes are essential for *M. tuberculosis* H37Rv to survive in the macrophage under conditions that model the immune response (Fig. 2B) (Rengarajan *et al*., 2005). Most of the 11 encode enzymes such as KSTD, HsaA and HsaD, which are involved in the degradation of steroid rings A and B (Fig. 1). Intriguingly, cholesterol catabolic genes that were not identified in TraSH studies have functions that may be complemented by other similar genes in *M. tuberculosis* H37Rv. These include KshB and HsaB, the respective reductase components of the AD(D) and the 3-HSA hydroxylases. Some of the TraSH mutants, such as *mce4*, displayed a progressive *in vivo* growth defect 2-4 weeks after infection in mice (Sassetti and Rubin, 2003). Moreover, the essentiality of some of these genes has been substantiated. Thus, a Δ*yrbE4A* (i.e. *supA*) and Δ*mce4* mutants show attenuated survival of *M. tuberculosis* H37Rv in mice (Sassetti and Rubin, 2003; Joshi *et al*., 2006). Similarly, inactivation *mt3626* of *M. tuberculosis* CDC1551 (*rv3527* in H37Rv, Fig. 2B), a gene of unknown function adjacent to *kshA* and clustered with the *hsa* genes, had an impaired ability to arrest phagosome acidification and resulted in attenuated survival (Pethe *et al*., 2004). Clearly, the essential nature of the cholesterol

catabolic genes needs to be further substantiated. However, the available evidence suggests that cholesterol uptake and metabolism are important for *M. tuberculosis* to be able to persist in the macrophage for longer periods of time.

The deduced cholesterol catabolic pathway is consistent with at least two features of *M. tuberculosis* pathogenicity. First, cholesterol is essential for the phagocytosis of the bacterium by the macrophage as well as for inhibition of phagosome maturation (de Chastellier and Thilo, 2006, Gatfield and Pieters, 2000, Peyron *et al*., 2000). For example, depletion of cholesterol from macrophages abrogates the receptor-specific uptake of mycobacteria. Moreover, cholesterol depletion overcomes the block in phagosome maturation of *M. avium*-infected macrophages (de Chastellier and Thilo, 2006), further suggesting that cholesterol might play a similar role in other mycobacterial pathogens. Second, the large number of oxygenases in the pathway is consistent with the observation that TB infections are associated with the most O_2-rich sites within the body (Adler and Rose, 1996). More specifically, the cholesterol catabolic genes encode six oxygenases including two associated cytochromes P450 of unknown function. Reactivation of the disease occurs most frequently in the upper pulmonary lobes, likely the most oxygenated regions of the body (Adler and Rose, 1996).

At least two differences between the deduced cholesterol catabolic pathways in RHA1 and the pathogenic mycobacteria suggest distinct metabolism of cholesterol rings C and D. First, the Baeyer-Villiger monooxygenase and hydrolase typically associated with the ring fission of cycloalkanones, and whose genes are up-regulated in the RHA1 cholesterol transcriptome, are not conserved in the mycobacteria. Second, the mycobacterial *hsa* operon includes an *N*-acetyl transferase gene (Anderton *et al*., 2006; Cole *et al*., 1998). Thus, it is possible that pathogenic mycobacteria transform this portion of the cholesterol molecule for an alternate function such as signaling or cell wall integrity. Moreover, the cholesterol metabolic enzymes reported herein may also transform other host steroids or their derivatives, such as vitamin D, recently shown to mediate an innate immune response to mycobacteria (Liu *et al*., 2006). Nevertheless, the identified mycobacterial pathway transforms most of the cholesterol molecule to central metabolites, consistent with growth of bacillus Calmette-Guérin on cholesterol *in vitro* and suggesting that the sterol is an important source of energy for *M. tuberculosis* during its survival in the macrophage. The essential nature of the cholesterol catabolic enzymes *in vivo* makes them promising new targets for the development of novel therapeutic agents to combat XDR-TB and other strains, particularly as many of these enzymes have no human homolog.

MATERIALS AND METHODS

Bacterial growth

RHA1 was grown at 30°C on a shaker in one of two minimal media: W minimal salt medium (Seto *et al*., 1995) plus 20 mM pyruvate or 2 mM cholesterol; or a similar medium supplemented with a different mineral solution (Fortin *et al*., 2005) plus either cholesterol, AD, pyruvate or benzoate as indicated. RHA1 cells were harvested at mid-log phase (OD_{600}

of 1.0 for pyruvate, and 2.0 for cholesterol). bacillus Calmette-Guérin was grown at 37°C on a tube roller (10 rpm) in screw-capped 15-ml vials filled with 10 ml liquid medium containing 0.5 g/ml asparagine, 1 g/ml KH_2 PO_4, 2.5 g/ml Na_2PO_4, 10 mg/l $MgSO_4.7H_2O$, 0.5 mg/l $CaCl_2$, 0.1 mg/l $ZnSO_4$, 50 mg/l ferric ammonium citrate, 0.5 ml/l Triton wR1339 (Tyloxapol) (18) plus the indicated amount of cholesterol, 10 mM pyruvate or 10 mM glucose. Total protein content of cultures was determined in cells disrupted by sonication (10 cycles of 30s at 6 micron) using the Bradford protein assay (BioRad) and bovine serum albumin as standard.

RNA extraction and microarray

RNA was isolated from RHA1 as described previously (Gonçalves *et al*., 2006). RNA was similarly isolated from bacillus Calmette-Guérin except that both the RNeasy Plus and RNeasy Mini Kits (Qiagen) were used and the sample was treated with 2 U TURBO DNase (Ambion, Austin, TX). The RHA1 transcriptome was analyzed by using indirectly labeled cDNA and a microarray containing 70-mer probes for 8,313 genes as described (Gonçalves *et al*., 2006). Data were analyzed by using GeneSpring (Agilent Technologies, Santa Clara, CA) and MeV 3.1 (The Institute for Genomic Research, Rockville, MD) For each condition, RNA was extracted from each of three independently grown cultures. Data were averaged and normalized using Locally Weighted Linear Regression (Lowess). Details of the microarray design, transcriptomic experimental design and transcriptomic data were deposited in the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE6709.

Quantitative RT-PCR

RT-PCR was performed as described previously (Gonçalves *et al*., 2006) using TaqMan probes, and cDNA was synthesized using the ThermoScript RT-PCR System (Invitrogen, Carlsbad, CA) and random hexamers. All oligonucleotide and probe sequences are provided in supplemental Table S3. The gene encoding DNA polymerase IV and σ^A were used as internal standards in the multiplex reactions performed using RHA1 and bacillus Calmette-Guérin cDNA, respectively (Gonçalves *et al*., 2006). The C_t values were normalized (ΔC_t) by subtracting those of the internal standard. Significant differences in ΔC_t values were tested using a two-sample t-test assuming unequal variances. Relative fold differences were calculated as $2^{-\Delta\Delta Ct}$, where $\Delta\Delta C_t = \Delta C_{t\ treatment} - \Delta C_{t\ control}$.

Gene replacement and deletion

The *hsaC* gene was replaced in RHA1 with an apramycin resistance marker, *apra*^{R}, using a procedure in which the gene was first replaced in a fosmid using a λ-RED-based methodology and then in RHA1 using the modified fosmid and allelic exchange (Patrauchan *et al*., 2005). The parent fosmid, RF00128O15, contained a 38.3 kb of RHA1 DNA including the *hsaADCB* cluster. The oligonucleotides used to generate the resistance cassette used to replace *hsaC* were: hsaC-for1 and hsaC-rev1 (supplemental Table S3). The 6 *mce4* genes and the *supAB* genes were deleted separately in RHA1 using the *sacB* counter-selection system

essentially as described (van der Geize *et al*., 2001). Oligonucleotides used to amplify the upstream and downstream region of the 6 *mce4* genes were ro04698-F and ro04698-R(*Spe*I), and ro04703-F(*Spe*I) and ro04703-R(*Hind*III), respectively. The upstream and downstream region of the *supAB* genes were amplified using oligonucleotides SupA-F and SupA-R(*Spe*I), and Sup4B-F(*Spe*I) and SupB-R (supplemental Table S3). Gene replacements and deletions were verified using a series of PCR reactions using: (a) primers matching sequences within the target gene(s), (b) primers matching sequences flanking the target gene, and when appropriate (c) primers matching a region within *apra*^{R}.

Cloning and expression of Mtb genes

The *hsaC*$_{H37Rv}$ and *hsaD*$_{H37Rv}$ genes were amplified by PCR using Expand High Fidelity^{TM} DNA polymerase (Roche Applied Sciences, Laval, P.Q., Canada) and cloned essentially as described for *dbfB* (Fortin *et al*., 2005). The genes were amplified using *M. tuberculosis* H37Rv genomic DNA and either Hcmt-F and Hcmt-R, or Hdmt-F and Hdmt-R (supplemental Table S3). The amplicons were digested with *Nde*I and *BamH*I, cloned into similarly digested pT7-7, and their respective nucleotide sequences confirmed to yield pT7HC1 and pT7HD1. HsaC$_{H37Rv}$ and HsaD$_{H37Rv}$ were produced using *E. coli* GJ1158 transformed with pT7HC1 and pT7HD1, respectively, as described for DbfB (Fortin *et al*., 2005).

Enzyme assays

HsaC$_{H37Rv}$ and HsaD$_{H37Rv}$ activities in cellular extracts were measured by following the formation (HsaC) or consumption (HsaD) of the ring-cleaved product on a Varian Cary 5000 spectrophotometer equipped with a thermojacketed cuvette holder, essentially as described for biphenyl catabolic enzymes (Fortin *et al*., 2005). Experiments were performed using 20 mM 3-[4-(2-hydroxyethyl)-1-piperazinyl]propanesulfonic acid, 80 mM sodium chloride, pH 8.0 at 25.0 ± 0.1 °C. Concentrations of 4,9-DSHA (ε_{392} = 7.64 $mM^{-1}cm^{-1}$) and 2-hydroxy-6-oxo-6-phenylpentadienoate (HOPDA; ε_{434} = 32.5 $mM^{-1}cm^{-1}$) were monitored at 392 nm and 434 nm, respectively. Initial velocities were determined from a least-squares analysis of the linear portion of the progress curves using the kinetics module of the Cary software. Steady-state rate equations were fit to data as described previously (Fortin *et al*., 2005).

Metabolite preparation and characterization

Culture supernatant was acidified using 0.5% orthophosphoric acid then extracted twice with 0.5 volume of ethyl acetate. The ethyl acetate extracts were pooled, dried using anhydrous magnesium sulfate and evaporated to dryness using a rotary evaporator. The residue was dissolved in a 7:3 mixture of methanol:water containing 0.5% phosphoric acid and purified by HPLC using a Waters 2695 separation module and a Prodigy 10-µm ODS-Prep column (21.2 x 250 mm; Phenomenex, Torrance, CA). Metabolites were eluted using the same methanol:water solvent at a flow rate of 5 ml/min. The eluate was monitored at 280 nm. The retention time of the major metabolite was ~21 minutes. The fractions

containing this metabolite were pooled, added to 10 volumes of water and extracted as described above. The metabolite was derivatized using Sylon BFT (Supelco, Inc.) and analyzed using an Agilent 6890 gas chromatograph and Agilent 5973N mass-selective detector in electron ionization mode. The extinction coefficient of 4,9-DHSA was determined using an oxygraph assay (Fortin *et al*., 2005).

ACKNOWLEDGEMENTS

We thank Pascal D. Fortin for help in cloning hsaC and hsaD; Yossef Av-Gay (University of British Columbia) for the gift of H37Rv genomic DNA; Christine Florizone, Gordon R. Stewart, Matthew J. Myhre, and Jie Liu for skilled technical assistance; and Charles Thompson for critically reading the manuscript. This work was supported by grants from Genome Canada (to L.D.E. and W.W.M.), GenomeBC(to L.D.E. and W.W.M.), the Natural Sciences and Engineering Research Council of Canada (to L.D.E. and W.W.M.), the Integration of Biosynthesis and Organic Synthesis Program of Advanced Chemical Technologies for Sustainability (to M.H.W.), and the Wellcome Trust (to E.S. and M.C.A.). K.Y. received a postgraduate scholarship from the Natural Sciences and Engineering Research Council of Canada. M.H.W. received a traineeship from the European Graduate College program of the Groningen Biomolecular Sciences and Biotechnology Institute.

REFERENCES

References are listed on pages 119-140.

CHAPTER 3

Cytochrome P450 125 (CYP125) catalyzes C26-hydroxylation to initiate sterol side chain degradation in *Rhodococcus jostii* RHA1

Kamila Z. Rosłoniec[#], Maarten H. Wilbrink[#], Jenna K. Capyk, William W. Mohn, Martin Ostendorf , Robert van der Geize, Lubbert Dijkhuizen and Lindsay D. Eltis
# These authors contributed equally to this work.

Published in Mol. Microbiol. (2009) 74: 1031-1043

ABSTRACT

The *cyp125* gene of *Rhodococcus jostii* RHA1 was previously found to be highly up-regulated during growth on cholesterol and the orthologue in *Mycobacterium tuberculosis (rv3545c)* has been implicated in pathogenesis. Here we show that *cyp125* is essential for *R. jostii* RHA1 to grow on 3-hydroxysterols such as cholesterol, but not on 3-oxo sterol derivatives, and that CYP125 performs an obligate first step in cholesterol degradation. The involvement of *cyp125* in sterol side chain degradation was confirmed by disrupting the homologous gene in *Rhodococcus rhodochrous* RG32, a strain that selectively degrades the cholesterol side chain. The RG32Ω*cyp125* mutant failed to transform the side chain of cholesterol, but degraded that of 5-cholestene-26-oic acid-3β-ol, a cholesterol catabolite. Spectral analysis revealed that while purified ferric $CYP125_{RHA1}$ was <10% in the low-spin state, cholesterol (K_D^{app} = 0.20 ± 0.08 μM), 5α-cholestanol (K_D^{app} = 0.15 ± 0.03 μM) and 4-cholestene-3-one (K_D^{app} = 0.20 $\pm$ 0.03 μM) further reduced the low spin character of the heme iron consistent with substrate binding. Our data indicate that CYP125 is involved in steroid C26-carboxylic acid formation, either catalyzing the oxidation of C26 to the corresponding carboxylic acid or to an intermediate state.

INTRODUCTION

Cytochromes P450 (P450s) are a widely distributed class of heme-containing monooxygenases that are present in all domains of life. Their essential roles in diverse metabolic pathways have also generated considerable interest for their use as biocatalysts (Julsing *et al*., 2008). Genome sequence data analysis has revealed that *Actinobacteria* posses a remarkable number of genes encoding P450s compared to other prokaryotes (McLean *et al*., 2006; Lamb *et al*., 2006). For example, *Rhodococcus jostii* RHA1 harbors 29 genes predicted to encode P450s (McLeod *et al*., 2006). While the biological function of most of these monooxygenases is still unknown, several of them have been implicated in sterol/steroid catabolism.

The microbial degradation of cholesterol (5-cholestene-3β-ol; Fig. 1, compound I) involves two processes: sterol side chain elimination and steroid ring opening (van der Geize and Dijkhuizen, 2004). The order of these two processes *in vivo* is unknown and may vary between microorganisms. Generally, oxidation of the cholesterol 3β-hydroxyl moiety and isomerization of Δ5 into Δ4 is thought to initiate sterol degradation (Sojo *et al*., 1997; Chen *et al*., 2006; Chiang *et al*., 2008). This transformation is catalyzed by either cholesterol oxidase (CHO, (MacLachlan *et al.*, 2000)) or 3β-hydroxysteroid dehydrogenase (3β-HSD, (Yang *et al*., 2007)) and results in the formation of 4-cholestene-3-one (Fig. 1, compound II). Further degradation of 4-cholestene-3-one proceeds via hydroxylation at C26 to initiate side chain degradation or oxidation of rings A and B analogous to ring degradation of 4-androstene-3,17-dione (AD), resulting in the formation of 2-hydroxyhexa-2,4-diene-oic acid (Fig. 1, compound VI; van der Geize *et al*., 2007). Microbial sterol side chain degradation has been studied at the biochemical level in more detail in *Nocardia* species and *Mycobacterium sp.* strains NRRL B-3683 and NRRL B-3805 (Sih *et al*., 1968a, 1968b; Marsheck *et al*., 1972; Fujimoto *et al*., 1982a, 1982b). The latter two are capable of selectively degrading the 17-alkyl side chains of cholesterol and phytosterols. Microbial cholesterol side chain degradation is initiated by C26 hydroxylation followed by further oxidation to the sterol C26-oic acid (Fig. 1, compound III). Subsequent degradation occurs via a mechanism similar to β-oxidation of fatty acids that leads to the formation of a steroid C22-oic acid intermediate (Fig. 1, compound IV) with the concomitant release of propionyl-CoA and acetyl-CoA. The remaining C3 side chain is released as propionyl-CoA via a different mechanism (Sih *et al*., 1967, 1968b).

Rhodococcus rhodochrous DSM43269 (synonym IFO3338) is able to selectively degrade the sterol side chain in the presence of iron chelators, which inhibit 3-ketosteroid 9α-hydroxylase (KSH) activity (Arima *et al*., 1978). This phenotype was replicated in a stable multiple gene deletion mutant strain of *R. rhodochrous* DSM43269 (strain RG32) lacking KSH activity (Wilbrink *et al.,* submitted, [Chapter 4]). Mutant strain RG32 is completely blocked in steroid ring degradation and capable of selective sterol side chain degradation, thereby accumulating 1,4-androstadiene-3,17-dione (ADD) (Fig. 1, compound V) and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC) (Fig. 1, compound IV) from sterols. The strain RG32 phenotype thus allows us to specifically analyze sterol side chain degradation.

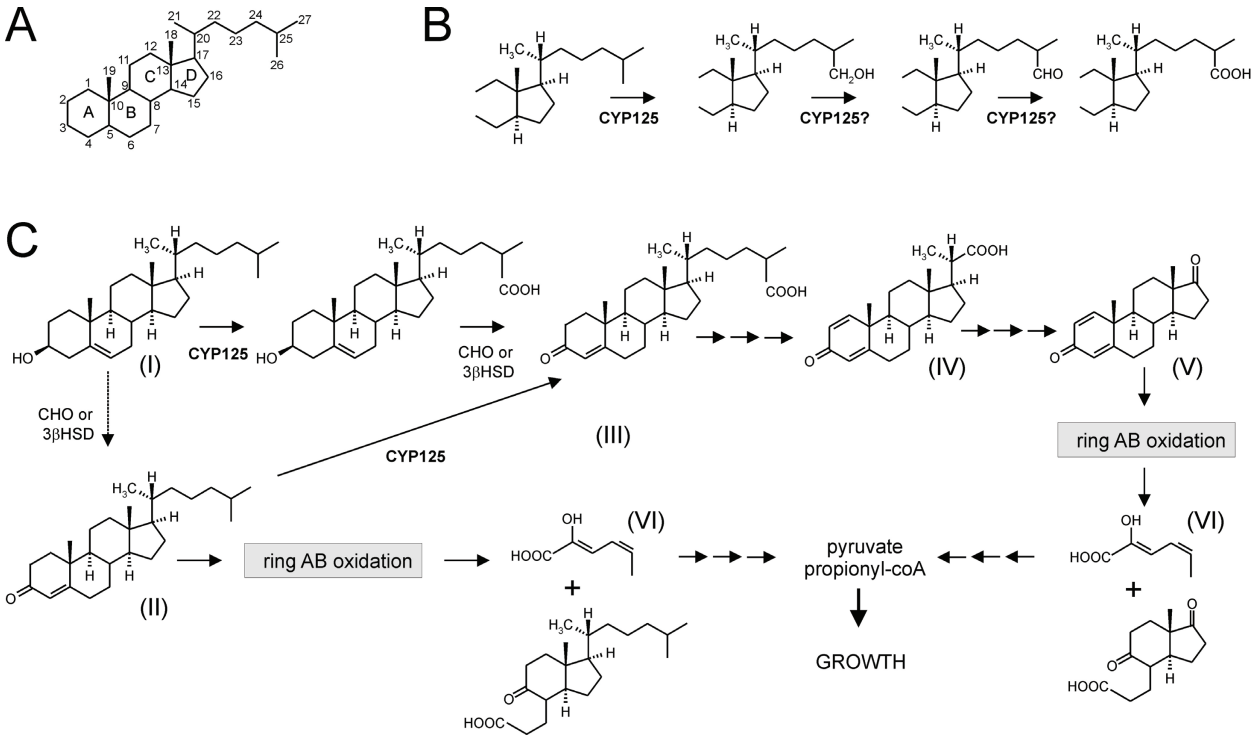


Fig. 1. The initial steps of aerobic cholesterol degradation in bacteria (Sih *et al*., 1968a, 1968b; Szentirmai, 1990; van der Geize *et al.,* 2007). (A) Steroid nomenclature. (B) CYP125 is involved in steroid C26 hydroxylation. Subsequent oxidation leads to a C26-oic acid metabolite. (C) Sterol degradation proceeds via steroid ring oxidation and side chain degradation (upper route). The exact order of side chain degradation and ring oxidation *in vivo* is unknown and may vary between microorganisms. In *R. jostii* RHA1, ring oxidation is not initiated until sometime after the side chain attack by CYP125 (dotted arrow). The depicted metabolites are: (I) 5-cholestene-3β-ol (cholesterol), (II) 4-cholestene-3-one, (III) 5-cholestene-26-oic acid-3β-ol, (IV) 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC), (V) 1,4-androstadiene-3,17-dione (ADD), and (VI**)** 2-hydroxyhexa-2,4-diene-oic acid. *R. rhodochrous* mutant strain RG32 (see text) converts compound I into compounds IV and V by selective side chain degradation. Abbreviations: CYP125, steroid 26-monooxygenase; CHO, cholesterol oxidase; 3βHSD, 3β-hydroxysteroid dehydrogenase (Yang *et al*., 2007).

To date, genes involved in sterol side chain degradation have not been identified. Using transcriptomic analysis, we recently identified a cholesterol catabolic gene cluster in *Rhodococcus jostii* RHA1 that includes two P450-encoding genes (van der Geize *et al*., 2007). Interestingly, *ro04679* (*cyp125*$_{RHA1}$) was one of the most highly up-regulated genes within this cluster during growth on cholesterol, suggesting an important role for this enzyme in cholesterol catabolism. In the RHA1 genome, *cyp125* is located proximal to genes predicted to encode β-oxidation enzymes, and suggested to be involved in degradation of the alkyl side chain of cholesterol (van der Geize *et al*., 2007). Moreover, *cyp125* is located within the *ro04482-ro04705* gene cluster encompassing the *mce4* genes, which encode the uptake system for cholesterol and related steroids with unsubstituted alkyl side chains (Mohn *et al*., 2008).

Here we report the molecular characterization of CYP125 as a steroid 26-monooxygenase. The *cyp125* gene was inactivated in each of *R. jostii* RHA1 and *R. rhodochrous* RG32 and the

effect on cholesterol catabolism was elucidated. $CYP125_{RHA1}$ was heterologously expressed and purified, and its binding to cholesterol and its analogues was investigated. This study provides novel insights into bacterial steroid degradation, revealing that degradation in *R. jostii* RHA1 is initiated by side chain oxidation, not oxidation of the rings.

RESULTS

CYP125 possesses conserved amino acid residues for interaction with sterols

Bioinformatic analysis revealed that $CYP125_{RHA1}$ has high amino acid sequence identity with P450s from other *Actinobacteria*, including *Nocardia farcinica* strain IFM10152 (Nfa5180, 79% (Ishikawa *et al*., 2004)) and *Mycobacterium tuberculosis* strain H37Rv (Rv3545c, 69% (Cole *et al*., 1998; Camus *et al.,* 2002)). These proteins belong to the uncharacterized CYP125 family (subfamily A) of P450 enzymes (Nelson *et al*., 1996), in which $CYP125_{RHA1}$ has been assigned CYP125A14P (http://drnelson.utmem.edu/biblioE.html#125). These monooxygenases presumably transform lipid-like compounds, as the CYP125 family includes many actinobacterial proteins associated with lipid degradation (Ventura *et al*., 2007).

Bioinformatic analysis further revealed that the annotated sequence of $CYP125_{RHA1}$ was about 50 residues longer than that of the annotated orthologues. Careful analysis of the $cyp125_{RHA1}$ nucleotide sequence indicated that the start codon most likely is located 159 nucleotides downstream from that in the original annotation, and is preceded by a Shine-Dalgarno sequence (aggag). Thus, $cyp125_{RHA1}$ is a gene of 1,257 nucleotides, encoding a protein of 418 amino acids with a calculated molecular mass of 47.2 kDa. The re-annotated sequence of $cyp125_{RHA1}$ (RHA1 genome coordinates 4930900...4932156) was used in this study.

Amino acid sequence alignments revealed that the actinobacterial CYP125s share the conserved motifs characteristic for the P450 super-family, as well as key residues of cholesterol-transforming eukaryotic P450s (supplemental Fig. S1). The latter belong to various families, including: CYP3A4, which performs 4β-hydroxylation of cholesterol; CYP11A1, which transforms cholesterol to pregnenolone via C20-C22 bond-cleavage; CYP27A1, which hydroxylates cholesterol at C27; and CYP46A1, which transforms cholesterol to 24*S*-hydroxycholesterol (Mast *et al.,* 2006; Pikuleva, 2006; Storbeck *et al.,* 2007). The presence of these conserved residues in CYPs125 and in P450 enzymes known to interact with sterols suggests that sterols are substrates for CYP125.

CYP125 is essential for growth on 3-hydroxy-sterols

To elucidate the role of *cyp125* in sterol/steroid catabolism, an unmarked single gene deletion mutant strain, RHA1Δ*cyp125*, was constructed. Growth experiments in mineral medium supplemented with cholesterol revealed that the RHA1Δ*cyp125* strain was unable to grow on cholesterol (Table 1). To confirm that the observed phenotype was solely due to inactivation of *cyp125*, a complementation experiment was performed in which $cyp125_{RHA1}$ was supplied in *trans.* The complemented strain, RHA1Δ*cyp125* pTip-QC1$cyp125_{RHA1}$, displayed a restored wild type growth phenotype in mineral medium supplemented with

cholesterol (Table 1). Wild-type RHA1 and RHA1Δ*cyp125* were subsequently grown in mineral liquid media on a range of other sterols, steroids and their metabolites as sole carbon and energy sources (supplemental Table S1). RHA1 grew readily on all tested compounds. By contrast, RHA1Δ*cyp125* failed to grow on epicholesterol, 5α-cholestanol and on the plant sterol mixture β-sitosterol/β-sitostanol/campesterol. Remarkably, growth of strain RHA1Δ*cyp125* on 3-ketone oxidized derivatives of two of these sterols, 4-cholestene-3-one and 5α-cholestane-3-one, was unimpaired, likely due to degradation of the steroid ring structure (supplemental Table S1, Fig. 1). We thus conclude that CYP125 is essential for 3-hydroxy-sterol degradation. The phenotype of RHA1Δ*cyp125* was investigated further by growing the mutant in mineral liquid media supplemented with cholesterol and an additional non-repressing carbon source (*i.e*., pyruvate or glycerol). In contrast to the wild type strain, RHA1Δ*cyp125* did not significantly transform cholesterol under these conditions (Fig. 2).

Table 1. Growth in mineral media supplemented with cholesterol (2.5 mM) as sole carbon and energy source of wild type strain RHA1, mutant strain RHA1Δ*cyp125*, complemented mutant strain RHA1Δ*cyp125* (RHA1Δ*cyp125*+pTip-QC1*cyp125*), and RHA1Δ*cyp125* mutant strain harboring null vector (RHA1Δ*cyp125*+pTip-QC1) after 10 days of growth. Non-inoculated mineral medium with cholesterol was included as a negative control. Values represent mean ± standard deviation (n = 3).

Strain	Protein content (mg/L)	Residual cholesterol (%)
RHA1	49 ± 6	56 ± 5
RHA1Δ*cyp125*	5 ± 3	112 ± 4
RHA1Δ*cyp125*+pTip-QC1*cyp125*	57 ± 4	55 ± 4
RHA1Δ*cyp125*+pTip-QC1	3 ± 3	116 ± 16
Control (medium + cholesterol)	0	100 ± 13

To further investigate the initial cholesterol-transforming enzymes of RHA1, we assayed pyruvate-grown cultures of wild type RHA1 and mutant strain RHA1Δ*cyp125* that had been induced with cholesterol for total 3β-hydroxysteroid oxidation activity (Yang *et al*., 2007). These studies comprised assays for extracellular and intracellular activities arising from CHO and 3β-HSD. When cholesterol was used as a substrate in these assays, no activity was detected in either supernatants or cell lysates of these cultures, consistent with the lack of transformation of cholesterol by RHA1Δ*cyp125*. By contrast, 3β-hydroxysteroid oxidation activity was detected in lysates of cholesterol-induced cells of RHA1 (0.27 $\mu M \cdot min^{-1} \cdot mg^{-1}$) and RHA1Δ*cyp125* (0.76 $\mu M \cdot min^{-1} \cdot mg^{-1}$) when 5-pregnene-3β-ol-20-one was used as a substrate in the assay. Overall, these data indicate that CYP125 is essential for cholesterol degradation by RHA1, and that it catalyzes an obligate first reaction in the cholesterol catabolic pathway.

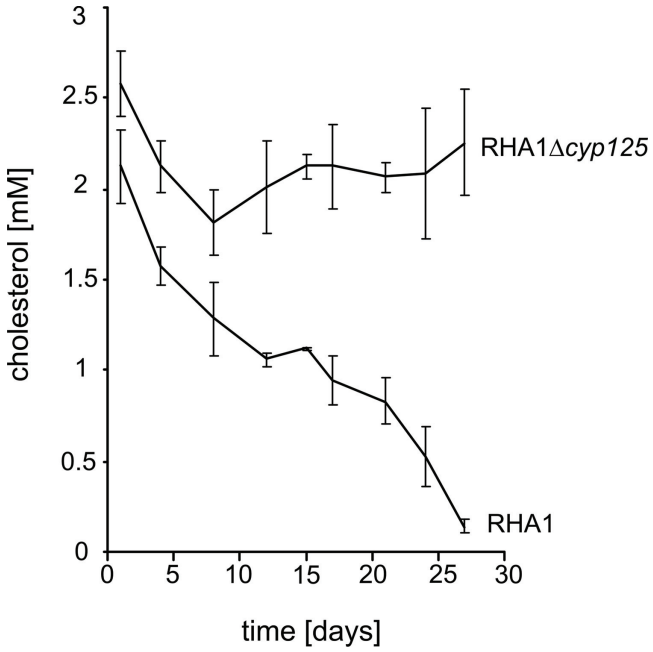


Fig. 2. Cholesterol degradation by cell cultures of strains RHA1 and RHA1Δ*cyp125* grown in mineral liquid media supplemented with pyruvate (20 mM) and cholesterol (2.5 mM). The data represent averages of triplicates. Error bars indicate standard deviations.

CYP125 has a role in sterol side chain degradation

We hypothesized that $CYP125_{RHA1}$ might have a specific role in sterol side chain degradation. To substantiate this hypothesis, we used *R. rhodochrous* RG32, a mutant of *R. rhodochrous* DSM43269 which only degrades the side chain of cholesterol, transforming it to ADD (Fig. 1, compound V) and 1,4-BNC (Fig. 1, compound IV) (Fig. 3A). First, we cloned *cyp125* from *R. rhodochrous* DSM43269 by screening a genomic library of this strain with degenerate PCR primers based on conserved amino acid sequences found in actinobacterial CYP125s. A positive clone, containing 8.7 kb of insert DNA, was obtained, sequenced and analyzed. The insert carried $cyp125_{DSM43269}$, encoding a protein sharing 76% amino acid sequence identity with $CYP125_{RHA1}$ (supplemental Fig. S1). Moreover, the *cyp125* locus is similarly organized in *R. jostii* RHA1 and *R. rhodochrous* DSM43269. More specifically, the genes immediately downstream of *cyp125* in DSM43269 encode proteins sharing 56%, 74% and 86% amino acid sequence identity to those encoded by *ro04676*, *ro04677* and *ro04678*, respectively, in RHA1. Upstream of $cyp125_{DSM43269}$, orthologues of *ro04654* (82% identity) and *ro04653* (82% identity) were located, as well as genes encoding hypothetical proteins that have no counterparts in RHA1. We then specifically disrupted *cyp125* in RG32, yielding mutant strain RG32Ω*cyp125*. Whole-cell biotransformations of 3-hydroxy-sterols by RG32Ω*cyp125* revealed that the mutant was blocked in the ability to degrade sterol side chains (Fig. 3). Cell cultures of RG32Ω*cyp125* incubated with cholesterol showed no formation of ADD or 1,4-BNC (Fig. 3B). Similar results were obtained when RG32Ω*cyp125* cell cultures were incubated with 5α-cholestanol and β-sitosterol (data not shown). Contrary to RHA1Δ*cyp125*, cholesterol was rapidly converted by RG32Ω*cyp125* to 4-cholestene-3-one and 1,4-cholestadiene-3-one, which accumulated in the medium (Fig. 3B inset) Indeed, cholesterol-induced cells of RG32 and RG32Ω*cyp125* contained high levels of 3β-hydroxysteroid total oxidation activity using cholesterol as a substrate (0.34 and 0.73 $\mu M \cdot min^{-1} \cdot mg^{-1}$, respectively). By contrast, no extracellular activity was detected in either strain.

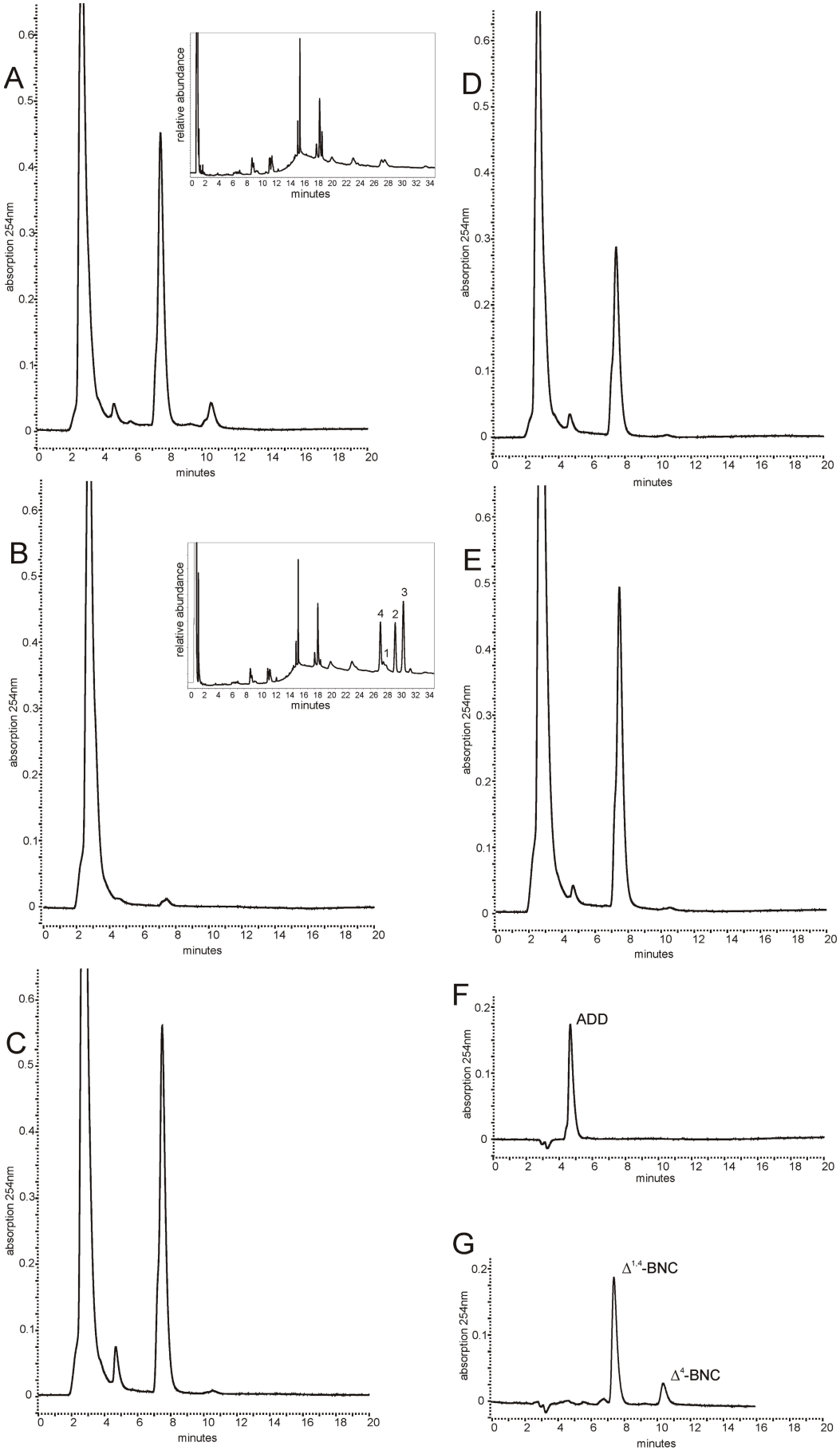
Reintroduction of *cyp125*$_{DSM43269}$ into RG32Ω*cyp125* under its native promoter fully restored the ability of the strain to degrade the cholesterol side chain (Fig. 3C). This excludes the possibility that side chain degradation in RG32Ω*cyp125* was blocked by polar effects rather than by disruption of *cyp125* directly.

CYP125$_{DSM43269}$ is involved in formation of the sterol C26-oic acid intermediate

We then tested the ability of mutant strain RG32Ω*cyp125* to convert each of two predicted sterol side chain degradation pathway intermediates: 5-cholestene-26-oic acid-3β-ol (Fig. 1, compound III) and the C24-oic intermediate 5-cholenic acid-3β-ol. Whole cell biotransformations performed with cultures of mutant strain RG32Ω*cyp125* resulted in conversion of both 5-cholenic acid-3β-ol and 5-cholestene-26-oic acid-3β-ol to ADD and 1,4-BNC (Fig. 3D, 3E). As predicted, RHA1Δ*cyp125* was able to grow on both of these compounds (supplemental Table S1). Both diastereomers of 5-cholestene-26-oic acid-3β-ol appeared to be metabolized, since 75 mol% of the added substrate was converted into ADD and 1,4-BNC. These results show that CYP125 is essential for the conversion of cholesterol into the C26-oic acid catabolite during sterol side chain degradation by both RG32 and RHA1.

Fig. 3. HPLC profiles of whole-cell biotransformations of cholesterol by cell cultures of (A) *R. rhodochrous* strain RG32 showing the formation of 1,4-androstadiene-3,17-dione (ADD) and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC) via selective sterol side chain degradation, (B) mutant strain RG32Ω*cyp125* and (C) *cyp125*$_{DSM43269}$ complemented mutant strain RG32Ω*cyp125*. HPLC profiles of whole-cell biotransformations of 5-cholenic acid-3β-ol (D) and 5-cholestene-26-oic acid-3β-ol (E) by cell cultures of *R. rhodochrous* mutant strain RG32Ω*cyp125* are also shown. Profiles of authentic ADD (50 µM, F) and 1,4-BNC (G), obtained by incubating authentic 3-oxo-23,24-bisnor-4-cholene-22-oic acid (50 µM, 4-BNC) with purified Δ1-KSTD1 (Knol *et al*., 2008), are included as reference samples.

Insets: GC profiles showing the accumulation of 4-cholestene-3-one (2), 1,4-cholestadiene-3-one (3) and 5α-cholestane-3-one (4) from cholesterol (1) by *R. rhodochrous* mutant strain RG32Ω*cyp125*, but not strain RG32.


A
B
C
D
E
F
G
absorption 254nm
relative abundance
minutes
1
2
3
4
ADD
Δ1,4-BNC
Δ4-BNC

Production and purification of $CYP125_{RHA1}$

To biochemically characterize $CYP125_{RHA1}$, we homologously produced and purified recombinant $CYP125_{RHA1}$ with a 6-histidine tag. Expression of *$cyp125_{RHA1}$* was first attempted in *E. coli* BL21(DE3) using T7 promoter-based expression vectors and conditions known to promote expression of P450 proteins, such as the addition of δ-aminolevulinic acid, $FeCl_3$, trace elements and thiamine (Parikh *et al*., 1997; Keizers *et al*., 2004). However, $CYP125_{RHA1}$ was not produced in significant amounts in *E. coli*. By contrast, *$cyp125_{RHA1}$* was well expressed in *R. jostii* RHA1 using the pTip-QC1 vector (Nakashima and Tamura, 2004). Addition of δ-aminolevulinic acid and other additives, usually necessary to promote expression of properly folded and soluble P450 proteins in *E. coli*, was not needed for homologous production of $CYP125_{RHA1}$ in *R. jostii* RHA1.

$CYP125_{RHA1}$ was purified using Ni^{2+}-NTA affinity chromatography and was determined by SDS-PAGE analysis to be in excess of 95% pure. The CO-difference spectrum of purified $CYP125_{RHA1}$ displayed a maximum at 451 nm (Fig. 4, inset), indicating the heme iron thiolate ligation remained intact throughout the protein's purification. The absorption spectrum of the purified ferric $CYP125_{RHA1}$ had a maximum at 392 nm and a shoulder at 422 nm (Fig. 4). Based on analysis of this spectrum (see Experimental procedures), the preparation is estimated to contain ~93% high spin state heme iron.

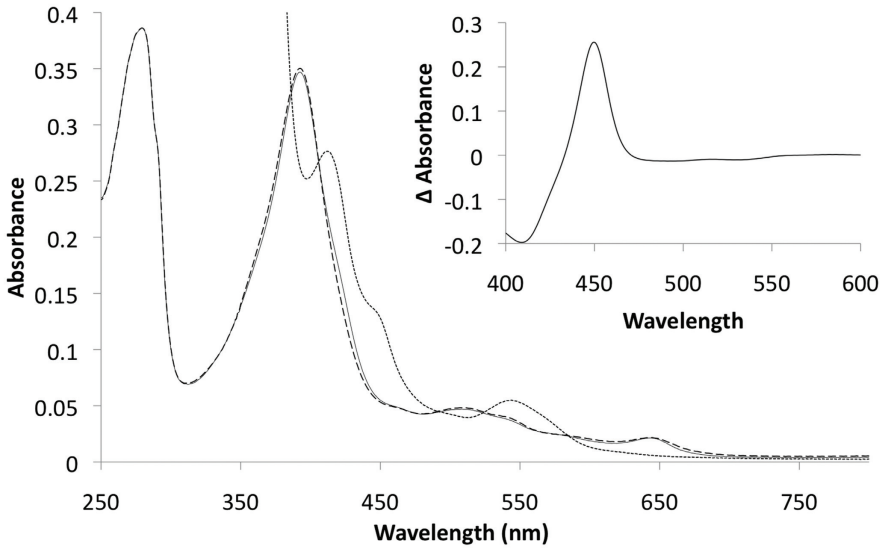


Fig. 4. The absorption spectrum of $CYP125_{RHA1}$ in the oxidized state as isolated (solid line) and incubated with 10 μM cholesterol in oxidized (dashed line) and reduced (dotted line) states. The inset shows the reduced CO-difference spectrum of the enzyme incubated with 10 μM cholesterol. The sample contained 2.9 μM purified $CYP125_{RHA1}$, 0.1 M potassium phosphate buffer, pH 7.0, 25°C; cholesterol was added from a 1 mM stock solubilized in 10% 2-hydroxypropyl-β-cyclodextrin.

Spectroscopic analysis of sterol binding

Spectroscopic assays were performed with purified $CYP125_{RHA1}$ to investigate its binding to sterols. Following the addition of cholesterol (Fig. 5A) or 5α-cholestanol (data not shown) in a solution of 10% 2-hydroxypropyl-β-cyclodextrin, $CYP125_{RHA1}$ exhibited a spectral change

with a pronounced trough at 422 nm and a peak at 392 nm, consistent with the decrease in the low-spin character of the heme iron associated with substrate binding. The difference spectrum also exhibited a perturbation at 395 nm in comparison to the typical type I binding spectrum. A perturbation at the same wavelength was observed upon addition of 5-cholestene-26-oic acid-3β-ol in 10% 2-hydroxypropyl-β-cyclodextrin, although the acid elicited no underlying type I spectral change at concentrations up to 20 μM (Fig 5A). Cholesterol also induced a type I binding spectrum when added in the presence of other solubilizing agents, such as Triton WR1339 and dimethylsulphoxide. However, the spectral shifts were much weaker than in the presence of 2-hydroxypropyl-β-cyclodextrin (data not shown).

Using equation 1, apparent K_D values for cholesterol, 5α-cholestane-3β-ol, and 4-cholestene-3-one were evaluated to be 0.20 ± 0.08 μM, 0.15 ± 0.03 μM, and 0.20 ± 0.03 μM respectively. The concentrations of enzyme calculated using this equation (4.0, 4.3, and 3.6

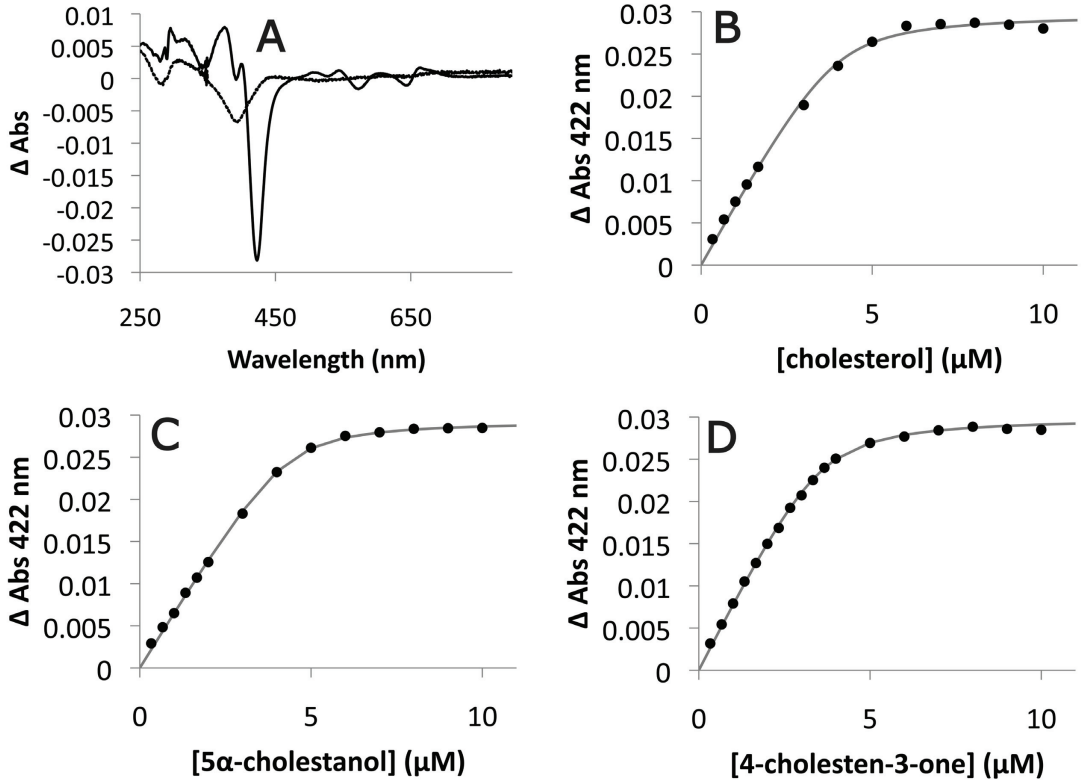


Fig. 5. Binding of steroids to purified $CYP125_{RHA1}$. (A) Spectral responses of 3.7 μM purified $CYP125_{RHA1}$ induced by 10 μM cholesterol (solid line) and 10 μM 5-cholestene-26-oic acid-3β-ol (dashed line). The dependence of the absorbance change of $CYP125_{RHA1}$ at 422 nm on (B) cholesterol, (C) 5α-cholestanol, and (D) 4-cholestene-3-one concentration. The best fit of equation 1 to the data as determined using R is represented as a grey line with fitted parameters K_D = 0.20 ± 0.08 μM, ΔA_{max} = 0.0298 ± 0.0006, and [E] = 4.0 ± 0.2 μM for cholesterol; K_D = 0.15 ± 0.03 μM, ΔA_{max} = 0.0293 ± 0.0002, and [E] = 4.3 ± 0.1 μM for 5α-cholestanol; and K_D = 0.20 ±0.03 μM, ΔA_{max} = 0.0300 ± 0.0002, and [E] = 3.6 ± 0.1 μM for 4-cholestene-3-one. Steroids were prepared as stock solutions in 10% 2-hydroxypropyl-β-cyclodextrin which alone did not induce a $CYP125_{RHA1}$ spectral response.

μM, respectively) were within 15% of the enzyme concentration calculated using the extinction coefficient for the reduced CO-difference spectrum of $\varepsilon_{450\text{-}490}$ = 91 $mM^{-1}cm^{-1}$ (3.7 μM), although this extinction coefficient has not been independently verified for this isozyme. The high quality fit of the equation to the binding data (Fig. 5B-D) supports a binding stoichiometry of 1:1 and suggests that $CYP125_{RHA1}$ does not harbor a ligand as isolated despite the proportion of high-spin iron. Finally, $CYP125_{RHA1}$ exhibited maxima at 451 nm in CO-difference spectra taken after each binding experiment, indicating that the heme-thiolate ligation remained intact.

DISCUSSION

The current study presents several lines of evidence identifying CYP125 as a steroid 26-monooxygenase that catalyzes the initial step in microbial sterol side chain degradation (Fig. 1). First, a *cyp125* deletion mutant of *R. jostii* RHA1 was unable to grow on or transform several 3-hydroxy-sterols with relatively long unactivated aliphatic side chains. Second, a *cyp125* disruption mutant of *R. rhodochrous* RG32 was completely blocked in cholesterol side chain degradation. However, this mutant was still able to degrade the side chain of 5-cholestene-26-oic acid-3β-ol (Fig. 1, compound III), an expected intermediate of cholesterol side chain degradation. Mutant RG32Ω*cyp125* thus is unable to form the sterol C26-oic acid intermediate, strongly indicating that $CYP125_{DSM43269}$ catalyzes the oxidation of the sterol at C26. Finally, $CYP125_{RHA1}$ bound cholesterol, 5α-cholestane-3β-ol, and 4-cholestene-3-one in a manner typical of P450 substrates: each compound induced a transition in the spin state of the heme iron and each bound with apparent submicromolar dissociation constants. The conclusion that CYP125 is a steroid 26-monooxygenase extends previous studies in which an NADH-dependent mixed function oxidase system was reported to be responsible for the first step in the mycobacterial sterol side chain degradation pathway (Szentirmai, 1990) catalyzing sterol C26-oxidation (Ambrus *et al*., 1995).

Our data indicate that the oxidation of C26 is an essential first step of cholesterol degradation in *R. jostii* RHA1. The RHA1Δ*cyp125* mutant not only failed to detectably transform cholesterol, but grew on 3-oxo steroids, such as 4-cholestene-3-one and 5α-cholestane-3-one, as effectively as the wild-type strain. This indicates that in *R. jostii* RHA1, C26-oxidation precedes oxidation of the 3β-hydroxyl moiety (Fig. 1). RHA1Δ*cyp125* is likely able to grow on 3-oxo steroids by degrading steroid A and B rings, resulting in the formation of 2-hydroxyhexa-2,4-diene-oic acid that is further metabolized to form pyruvate and propionyl-CoA for growth (Fig. 1; van der Geize *et al*., 2007). Previously, it was suggested that the microbial catabolism of cholesterol was initiated by ring oxidation (Sojo *et al*., 1997; Chen *et al*., 2006; Chiang *et al*., 2008). Indeed, *R. rhodochrous* RG32Ω*cyp125* is capable of performing ring oxidation in the absence of CYP125, illustrating that the order of ring oxidation and sterol side chain oxidation may vary between different species of bacteria. Consistent with the conclusion that $CYP125_{RHA1}$ initiates cholesterol degradation, genes encoding putative CHOs in RHA1 *(ro03863, ro04305*, *ro06201*) were not up-regulated during growth on cholesterol and are located outside of the cholesterol catabolic gene cluster

(McLeod *et al*., 2006; van der Geize *et al*., 2007). Although 3β-HSD has not been definitively identified in RHA1, *ro04707* encodes a protein sharing 43% amino acid similarity with 3β-HSD of *M. tuberculosis* (Rv1106c) and is located proximal to the genes encoding the Mce4 steroid transporter (Mohn *et al.,* 2008). Indeed, *ro04707* was up-regulated in cholesterol-grown RHA1 cells (van der Geize *et al*., 2007). While no cholesterol-transforming 3β-HSD activity was detected in RHA1, a 3β-HSD was expressed that transformed 5-pregnene-3β-ol-20-one, a 3β-hydroxysteroid with a shortened C21 side chain. $3\beta\text{-HSD}_{RHA1}$ thus appears to have a high substrate specificity for side chain-degraded cholesterol metabolites. This is similar to 3β-HSD of *M. tuberculosis* (Rv1106c) which had 3-fold higher activity towards 5-pregnene-3β-ol-20-one compared to cholesterol (Yang *et al.,* 2007). It is possible that in RHA1, side chain and ring degradation occur concurrently after C26 and C3-ol have been oxidized.

It is unclear whether CYP125 catalyzes the oxidation of C26 to the corresponding carboxylic acid or only to an intermediate state. Various P450s have been reported to catalyze multi-step oxidations (Helliwell *et al*., 1999, 2001; Ro *et al*., 2006), including a P450 from *Pseudomonas putida* PpG777 which catalyzes two sequential oxidations of linalool, to 8-hydroxylinalool and 8-oxolinalool, respectively (Ropp *et al*., 1993). It was proposed that a second oxygenation step results in a transient *gem*-diol adduct that spontaneously dehydrates to a more stable carbonyl compound (Ullah *et al*., 1990). Interestingly, $CYP125_{RHA1}$ displays significant amino acid sequence identity (32%) with linalool 8-monooxygenase and thus might well catalyze the complete oxidation of the aliphatic sterol side chain into the sterol 26-oic acid intermediate via a similar mechanism. Attempts to reconstitute the activity of $CYP125_{RHA1}$ have so far proved unsuccessful despite using a variety of electron donors, including the spinach ferredoxin and ferredoxin-reductase electron transport chain and the peroxide shunt using cumene hydroxyperoxide (Hrycay *et al.,* 1975). The physiological reductase of $CYP125_{RHA1}$ has not been identified yet.

M. tuberculosis contains a CYP125 encoded by *rv3545c* located within the recently described *igr* operon (Chang *et al.,* 2007; 2009). The bioinformatic data strongly suggest that the CYP125s of RHA1 and *M. tuberculosis* perform the same function: they are reciprocal best hits with 69% amino acid sequence identity that both occur in the cholesterol catabolic gene cluster (van der Geize *et al*., 2007). However, the recently reported phenotype of an Δ*igr* mutant indicates that Rv3545c is not a steroid 26-hydroxylase (Chang *et al.*, 2009): the mutant appeared to partially degrade cholesterol and transform the cholesterol labeled with ^{14}C at C26 into mycobacterial lipids. Additional studies are clearly required to definitively establish the role of CYP125 in *M. tuberculosis*. Indeed, while it is unclear if cholesterol degradation in mycobacteria occurs in the same manner as in *R. jostii* RHA1, two studies suggest that it does. First, *Mycobacterium sp*. NRRL B-3683, a mutant strain blocked in steroid ring degradation and able to selectively degrade the sterol side chain, displayed a clear preference for substrates possessing a 3β-hydroxy-Δ5 ring configuration compared to the 3-keto-Δ4 configuration (Marsheck *et al*., 1972). Second, 3β-HSD of *M. tuberculosis* (Rv1106c) showed 3-fold higher activity towards 5-pregnene-3β-ol-20-one, a sterol with a

C21 side chain, compared to cholesterol, suggesting that sterols with shortened side chains are preferred substrates of 3β-HSD (Yang *et al.,* 2007). Regardless of the precise function of CYP125 in *M. tuberculosis*, its gene is up-regulated during growth of *M. tuberculosis* in macrophages (Kendall *et al*., 2004) and $CYP125_{H37Rv}$ is more resistant to nitric oxide than other P450s of H37Rv (Ouellet *et al.,* 2009). Moreover, the gene appears to be important for infection in mice (Chang *et al*., 2007, 2009). CYP125 may thus be an interesting target for the development of novel anti-tuberculosis drugs.

MATERIALS AND METHODS

Bacterial strains, plasmids and chemicals

Plasmids and bacterial strains used are listed in supplemental Table S2. 5-Cholestene-3β-ol, 5α-cholestane-3β-ol, 4-cholestene-3-one and 5-cholestene-24β-ethyl-3β-ol (75%) were obtained from Sigma-Aldrich. 5α-Cholestane-3-one was obtained from Acros Organics. 5-Pregnene-3β-ol-20-one was obtained from ICN Biomedicals. 5-Cholestene-3α-ol, 23,24-bisnor-5-cholene-22-oic acid-3β-ol, 5-cholenic acid 3β-ol and 1-(5α)-androstene-3,17-dione were obtained from Steraloids. 4-Androstene-3,17-dione and 9,17-dioxo-1,2,3,4,10, 19-hexanorandrostane-5-oic acid were provided by Schering-Plough (Oss, The Netherlands).

Construction of R. jostii RHA1Δcyp125

A *cyp125* unmarked single gene deletion mutant of *R. jostii* RHA1 was constructed using the *sacB* counter-selection system (van der Geize *et al*., 2001). Genomic DNA of *R. jostii* RHA1 was isolated as described (van der Geize *et al*., 2000). Mutagenic plasmid pDEL$cyp125_{RHA1}$ was constructed for *cyp125* deletion, as follows. The upstream region of *cyp125* was amplified by PCR using forward primer 5' tcgacatccacttgatgaaggagaccg 3' and reverse primer 5'gcgACTAGTcactgctgtctcctgccctaagc 3', containing a *Spe*I restriction site (shown in capital letters). The resulting 1,421 bp amplicon was cloned into *Sma*I digested pK18*mobsacB*, resulting in pK18*mobsacB*UP*cyp125*. A 1,451 bp amplicon of the *cyp125* downstream flanking region including the *cyp125* stop codon was obtained using forward primer 5' cgcACTAGTtgacccctgattcagcggcgtcgg 3' (*Spe*I restriction site) and reverse primer 5' cgcAAGCTTgaacgaggacggcaagatcacgtccc 3' (*Hind*III restriction site). This amplicon was digested with *Spe*I/*Hind*III and ligated into *Spe*I/*Hind*III linearized pK18*mobsacB*UP*cyp125*, resulting in pDEL$cyp125_{RHA1}$. Deletion of *cyp125* from RHA1 was confirmed by PCR using forward 5' gcctcgacgattactggtgtgc 3' and reverse primer 5' cctcggacagaaggagaacagc 3'.

Functional complementation of mutant strain RHA1Δ*cyp125* was performed by electrotransformation (van der Geize *et al*., 2000) of RHA1Δ*cyp125* cells with expression plasmid pTip-QC1$cyp125_{RHA1}$ (see below).

Growth of R. jostii RHA1 and mutant RHA1Δcyp125 strain on sterols/steroids

Pre-cultures of wild type strain RHA1 and mutant strain RHA1Δ*cyp125* were grown for 3 days at 30°C with shaking (220 rpm) in mineral medium (MM, (Masai *et al*., 1995)) supplemented with pyruvate (20 mM) and used to inoculate MM liquid media (1:50)

supplemented with various sterols/steroids (1 g/l; supplemental Table S1) as sole carbon and energy source. Biomass production of *R. jostii* RHA1 cell cultures incubated with cholesterol were quantified by total protein content determination of sonicated cells (10 cycles of 30 s at 8 µm) using the Bradford protein assay (BioRad, Hercules, CA) with BSA as protein standard.

Biotransformation of cholesterol by R. jostii RHA1 and mutant RHA1Δcyp125

For biotransformation of cholesterol, pre-cultures of RHA1 and RHA1Δ*cyp125* were grown in MM supplemented with pyruvate (20 mM) for 3 days at 30°C with shaking (220 rpm). The pre-cultures were used to inoculate MM liquid media (1:50) containing pyruvate (20 mM) and cholesterol (2.5 mM).

Determination of intracellular and extracellular total 3β-hydroxysteroid oxidation activity

Total 3β-hydroxysteroid oxidation activity was determined by HPLC analysis essentially as described by Yang *et al*. (2007). Cell cultures of RHA1 and RHA1Δ*cyp125* were grown in MM supplemented with pyruvate (20 mM) for 3 days to an OD_{600nm} of 3. Cell cultures of RG32 and RG32Ω*cyp125* were grown overnight in LB medium. Grown cultures were induced for 16 hours by adding 0.5 mM cholesterol from a 10 mM stock prepared in isopropanol. The cell cultures (50 ml) were pelleted and the resulting supernatants were filter-sterilized and used for assaying extracellular cholesterol oxidation. The cell pellets were washed two times with 50 mM phosphate buffer (pH 7) supplemented with 5% (v/v) isopropanol and resuspended in 2 ml of the same buffer. Cell lysates were prepared by bead-beating. Cell lysates were centrifuged to remove cell debris. The 3β-hydroxysteroid oxidation assay was performed in 100 mM triethanolamide hydrochloride buffer (pH 8.5) supplemented with 0.05% (v/v) triton X-100, 3.5 mM NAD^+ and either 200 µM cholesterol or 200 µM 5-pregnene-3β-ol-20-one and incubated at 30°C for several hours (Yang *et al.,* 2007). 4-Cholestene-3-one and 4-progestene-3-one formation was quantified by HPLC-UV_{254nm} using calibration curves.

Steroid analysis

Steroid content of the cell cultures was analyzed by high-performance liquid chromatography (HPLC) and gas chromatography (GC). Culture samples (0.5 ml) were mixed with 2 ml of 80% methanol in water solution and filtered (0.2 µm) prior to analysis by HPLC-UV_{254nm}. HPLC was performed on an Alltima C18 column (250 x 4.6 mm; Alltech, Deerfield, USA, 35°C) using a mobile phase consisting of methanol:water (80:20) supplemented with 1% formic acid at a flow rate of 1 ml/min. For analysis of 4-cholestene-3-one and 1,4-cholestadiene-3-one a mobile phase consisting of acetonitrile:tetrahydrofuran (75:25) at a flow rate of 2 ml/min was used. Samples (0.5 ml) for GC analysis were mixed with 10% H_2SO_4 (10 µl) and ethyl acetate (2 ml) and the upper organic layer was subjected to GC. GC was performed on a (5% phenylo)-95% methoxypoly-siloxane Heliflex AT-5 ms column (30 m x 0.25 mm, ID x 0.25 µm; Alltech, Deerfield, USA) with FID-40 detection at 300°C.

Production of CYP125$_{RHA1}$

The *cyp125$_{RHA1}$* gene was amplified by PCR on genomic DNA of RHA1 with forward primer 5' CATATGgcgcagcccaatcttccag-aggg 3', containing an *Nde*I restriction site, and reverse primer 5' GGATCCtcagtgtctgaccgggcaaccg 3', containing a *Bam*HI restriction site, such that the recombinant protein contains a 6-histidine tag. PCR was performed in a reaction mixture (25 µl) consisting of Tris-HCl (10 mM, pH 8), polymerase buffer, dNTP (0.2 mM), primers (0.8 µM) and Vent polymerase (0.1 U, New England Biolabs, Ipswich, MA) under the following conditions: 5 min 95°C, 30 cycles of 45 s 95°C, 45 s 65°C, 2 min 72°C, followed by 5 min at 72 °C. A band of the expected size for *cyp125$_{RHA1}$* (1,266 bp) was purified from agarose gel using GenElute Gel Extraction Kit (Sigma-Aldrich, Steinheim, Germany) and cloned into *Sma*I digested pBlueScript KS(II) (Stratagene, La Jolla, CA, USA). The resulting plasmid was digested with *Nde*I and *Bam*HI and the DNA fragment containing *cyp125$_{RHA1}$* was ligated into *Nde*I/*Bam*HI digested pTip-QC1.

CYP125$_{RHA1}$ was homologously produced in *R. jostii* RHA1 using expression plasmid pTip-QC1*cyp125$_{RHA1}$*. Cells were cultured in Luria-Bertani broth (LB) in the presence of 25 µg/ml chloramphenicol. *R. jostii* RHA1 cells were transformed with pTip-QC1*cyp125$_{RHA1}$* by electroporation and grown on LB-agar plates containing 25 µg/ml chloramphenicol for two days, after which a single colony was used to inoculate 50 ml of liquid medium which was incubated at 30°C (200 rpm). When OD_{600} reached ~1.0 (~2-3 days) 2 l of medium inoculated with 20 ml of this pre-culture was incubated at 30°C. When the culture reached an OD_{600} of 0.6, thiostrepton was added to a final concentration of 50 µg/ml and the cells were incubated for a further 20 hrs before harvesting by centrifugation (4600 g, 4°C, 10 min) and subsequent washing with 0.1 M potassium phosphate buffer, pH 8.0. Cell pellets were flash frozen in liquid nitrogen and stored at -80°C until use.

Purification of CYP125$_{RHA1}$

The cell pellets were suspended in potassium phosphate buffer (pH 7.4) (Lussenburg *et al.*, 2005) containing DNase I (Roche diagnostics, IN). Cells were disrupted by bead beating and debris was removed by centrifugation at 10,000 x g for 45 min at 4°C. The clear supernatant was passed through a syringe-driven 0.45 µm filter. Cell free extracts were loaded on a NTA column (Qiagen) equilibrated with 0.1 mM potassium phosphate, pH 7.4. The protein was washed with Buffer A containing 0.5 M NaCl and a brown fraction eluted with buffer A further supplemented with 50 mM L-histidine. The protein was exchanged into 0.1 M potassium phosphate, pH 7.4, concentrated to 20 mg/ml, flash frozen as beads in liquid nitrogen and stored at -80°C. P450 protein concentrations were calculated from the reduced CO-bound difference spectrum using the extinction coefficient $\varepsilon_{450\text{-}490}$ = 91 $mM^{-1}cm^{-1}$ (Omura and Sato, 1964).

Spectroscopic analysis

UV-vis absorption spectra were recorded using a Cary 5000 spectrophotometer equipped with a thermojacketed cuvette holder (Varian, Walnut Creek, CA). The CO-bound form of $CYP125_{RHA1}$ was generated by first incubating samples with ~8 mM sodium dithionite for 10 min then slowly bubbling them with CO for 30 s. The proportion of purified protein containing high-spin ferric heme iron was estimated by comparing the spectra of $CYP125_{RHA1}$ to linear combinations of the spectra of $CYP125_{RHA1}$ in high and low spin states (Jung *et al*., 1991; Jefcoate, 1978) generated by adding 0.5% Triton X-100 and 40% methanol respectively to the sample. The same values were obtained when using substrate-free cytochrome $P450_{cam}$ from *Pseudomonas putida* as a low spin standard. Substrate-induced spectral responses were recorded in 0.1 mM KPi, pH 7.0 by titrating solutions of $CYP125_{RHA1}$ with 1.0 mM stock solutions of cholesterol, 5α-cholestane-3β-ol, and 4-cholesten-3-one in 10% 2-hydroxypropyl-β-cyclodextrin (Sigma). Equilibrium dissociation constants for were calculated using equation 1.

$$\Delta A = \frac{[S]_T + [E]_T + K_D - \sqrt{([S]_T + [E]_T + K_D)^2 - 4[S]_T[E]_T}}{2[E]_T} \Delta A_{Max} \quad (1)$$

In this equation, ΔA is the change in absorbance observed in the sample, $[S]_T$ is the total ligand concentration, $[E]_T$ is the total enzyme concentration, K_D is the equilibrium dissociation constant, and ΔA_{Max} is the change in absorbance at infinite ligand concentration. A non-linear least-squares fit of the equation to the data was obtained using the program R (http://www.R-project.org).

Construction of Rhodococcus rhodochrous RG32Ωcyp125

The *cyp125* orthologue in *R. rhodochrous* DSM43269 ($cyp125_{DSM43269}$) was identified using degenerate *cyp125* primers (forward 5' (a/g)ac(a/c/g/t)gc(a/c/g/t)cc(a/c/g/t)at(a/c/t) tggtg-gaa 3' and reverse 5' gg(a/g)tt(c/t)tc(a/g)aa(a/c/g/t) gc(a/g)tc(c/t)tc(a/g) 3') based on the deduced amino acid sequences T^{33}APIWWN^{39} and D^{329}EDAFENP^{336} from $CYP125_{RHA1}$; these sequences are highly conserved in Nfa5180 and Rv3545c from *N. farcinica* IFM10152 and *M. tuberculosis* H37Rv, respectively. A genomic library of *R. rhodochrous* DSM43269 in pRESQ (Petrusma *et al.,* 2009; supplemental Table S2) was screened by PCR using these degenerate primers. A single clone (pRESQ4679) was identified containing an 8.7 kb DNA insert. Nucleotide sequencing confirmed the presence of full-length (1,254 bp) *cyp125*.

The *cyp125* gene was disrupted in *R. rhodochrous* strain RG32 essentially as described (van der Geize *et al*., 2000). An internal $cyp125_{DSM43269}$ fragment of 811 bp was amplified by PCR using forward primer 5' gcacgaggaggtccgtgaggtc 3' and reverse primer 5' cgttgttggccgaggcgtacag 3' and ligated into *Eco*RV digested pK18*mobsac*B, yielding pΩ*cyp125* . This construct was used to transform *E. coli* S17-1 and was subsequently mobilized to mutant strain RG32 by conjugational transfer (van der Geize *et al*., 2001). Transconjugants were checked by PCR to confirm the *cyp125* gene disruption using forward primer 5' acgcagccaccgatgacctgtt 3', annealing to a sequence upstream of $cyp125_{DSM43269}$, and reverse primer 5' ctgcgtgcaatccatcttgttc 3', which is reverse complementary to part of the *aphII*

gene of pK18*mobsacB*. A PCR product of the expected size (1,903 bp) confirmed insertion of the disruption plasmid pΩ*cyp125* at the correct genomic locus.

Functional complementation RG32Ωcyp125

The intact *cyp125* $_{DSM43269}$ gene and its flanking regions were isolated from *Dra*III/*Bsp*HI digested pRESQ4679. A DNA fragment of 2.3 kb harboring *cyp125* was treated with T4 DNA polymerase and blunt-ligated into *Eco*RV digested shuttle vector pRRE1 (see below), resulting in pCOMP*cyp125*$_{DSM43269}$ that was used to transform electro-competent cells of RG32Ω*cyp125* as described (Fernandes *et al*., 2001). *E. coli-Rhodococcus* shuttle vector pRRE1 was constructed as follows. The *repA* and *repB* genes from *R. rhodochrous* DSM43269 endogenous plasmid pRC4 (GenBank/EMBL/DDJB accession number AB040101) were amplified from genomic DNA of strain DSM43269 using forward primer 5' cgatggcaagccaccgcgaagc 3' and reverse primer 5' atcggacagaagctgactaagg 3'. This amplicon (2.5 kb) was ligated into *Sma*I digested pK18mobsacB. A 2.6 kb *Eco*RI/*Xba*I DNA fragment of the latter construct was subsequently treated with Klenow fragment and blunt-ligated into *Psi*I digested pBs-Apra-ori (van der Geize *et al*., 2008a), resulting in pRRE1.

Whole-cell steroid biotransformations with RG32 and RG32Ωcyp125

Cell cultures of parent strain *R. rhodochrous* RG32, mutant strain RG32Ω*cyp125* and the *cyp125*$_{DSM43269}$ complemented mutant strain were grown overnight in liquid LB medium, supplemented with kanamycin 25 µg/ml when appropriate, at 30°C with shaking (200 rpm) until OD_{600} ~ 4 was reached. Sterols were added to the cell cultures at a final concentration of 0.5 mM from a 25 mM stock solution dissolved in acetone. Bioconversions were followed for 3 days of incubation at 30 °C with shaking (200 rpm). Accumulation of ADD and 1,4-BNC was analyzed by HPLC-UV_{254nm} as described above in "Steroid Analysis".

Chemical synthesis of 5-cholestene-26-oic acid-3β-ol

Synthesis of 5-cholestene-26-oic acid-3β-ol was carried out using a modification of the method described by Williams *et al.* (2002) with diosgenin as starting material (supplemental Fig. S2). In the first step, the 3-hydroxy group was protected as a methyl ether, using NaH and MeI and a reaction time of 24 h. The resulting 3-methyl ether (product 1) was isolated in near 100% yield after precipitation from water. Next, the ether rings were reductively ring-opened under Clemmensen conditions by treatment with Zn/HCl in ethanol at reflux temperature. After removal of the salts, extractive work up and a precipitation from acetone/water, the 16,27-dihydroxylated product (product 2) was obtained in near 100% yield. A region-selective protection of the primary alcohol at C27 was carried out by reaction with *tert*-butyldimethylsilyl chloride (TBDMSCl) and imidazole in DMF with 97% yield (product 3). For removal of the 16-hydroxy group, the Barton deoxygenation conditions were chosen. The C16-hydroxy group was transformed in the corresponding thiocarbonate with CS_2 under the influence of NaH. The intermediate thiocarbonate anion was quenched with methyl iodide. Next, a radical reduction reaction was carried out using

Bu_3SnH and AIBN. After purification by silica gel column chromatography, the TBDMS-protected 3-methyl ether form of 27-hydroxycholesterol (product 4) could be isolated in near 100% yield with an estimated ^{1}H-NMR purity of >80%. The TBDMS ether was removed under standard conditions using tetrabutylammonium fluoride (TBAF) in THF, and silica gel column chromatography was used to purify product 5 with a yield of 73%. Oxidation to the 26-oic acid was carried out under Jones' condition, using a mixture of sulfuric acid and chromine trioxide. 5-Cholestene-26-oic acid-3β-ol-3-methyl ether (product 6) was obtained by column chromatography purification in 89% yield with an estimated ^{1}H-NMR purity of 80%. The final step in the synthesis was the removal of the 3-methyl ether by treatment with TFA in DCM at room temperature for 2 days. After aqueous work up, the trifluoroethanol ester was saponified with K_2CO_3 in methanol and purified by silica gel column chromatography, generating 5-cholestene-26-oic acid-3β-ol (product 7) in a low yield of 15% with a ^{1}H NMR purity of approximately 95% and consisting of a 1:1 mixture of diastereomers at C26. Apparently, during the strong acidic conditions used for the removal of the 3-methyl ether, enolization and protonation at C26 had occurred giving rise to a 1:1 mixture of stereo-isomers. The structure was confirmed by mass spectrometry (supplemental Fig. S3).

ACKNOWLEDGEMENTS

We are grateful to N. P. E. Vermeulen, J. Commandeur and B. van Vugt-Lussenburg (VU Amsterdam) for technical advice and support. We thank T. Tiemersma-Wegman (University of Groningen) for technical assistance in GC analyses. We thank T. Tamura for kindly providing pTip-QC1. This work was supported by grants from the Integration of Biosynthesis and Organic Synthesis program of Advanced Chemical Technologies for Sustainability and the Canadian Institutes for Health Research (to L.D.E.). J.C. is the recipient of studentships from NSERC of Canada and MSFHR. We are grateful to N. P. E. Vermeulen, J. Commandeur and B. van Vugt-Lussenburg (VU Amsterdam) for technical advice and support. We thank T. Tiemersma-Wegman (University of Groningen) for technical assistance in GC analyses. We thank T. Tamura for kindly providing pTip-QC1.

REFERENCES

References are listed on pages 119-140.

CHAPTER 4

FadD19 of *Rhodococcus rhodochrous* DSM43269: a steroid-CoA ligase essential for degradation of C24-branched sterol side chains

Maarten H. Wilbrink, Lubbert Dijkhuizen and Robert van der Geize

Submitted for publication in Appl. Envir. Microbiol.

ABSTRACT

The cholesterol catabolic gene cluster of *Rhodococcus jostii* RHA1 contains a subset of genes that encode β-oxidation enzymes with a putative role in sterol side chain degradation. To further elucidate their *in vivo* physiological role, several homologous genes in *Rhodococcus rhodochrous* DSM43269, i.e. *fadD17, fadD19*, *fadE26*, *fadE27* and *ro04690*$_{DSM43269}$, were inactivated by mutagenesis in strain RG32, a five-fold *kshA* mutant of strain DSM43269. Strain RG32 lacks all 3-ketosteroid 9α-hydroxylase (KSH) activity, is fully blocked in steroid ring attack, but capable of selective sterol side chain degradation, accumulating 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid and 1,4-androstadiene-3,17-dione. Except for RG32Δ*fadD19*, none of the mutants revealed an aberrant phenotype on sterol side chain degradation compared to parent strain RG32. Unmarked deletion of *fadD19* in strain RG32 however completely blocked side chain degradation of the C24-branched sterols β-sitosterol and campesterol, but interestingly not that of cholesterol. The additional inactivation of *fadD17* in mutant RG32Δ*fadD19* also did not affect cholesterol side chain degradation. Heterologously expressed FadD19$_{DSM43269}$ nevertheless was active towards steroid-C26-oic acid substrates. Our data identified *fadD19* as a gene encoding steroid-CoA ligase activity with an essential *in vivo* role in the degradation of the side chains of C24-branched chain sterols. This is the first report on the identification and characterization of a CoA ligase with an *in vivo* role in sterol side chain degradation. The high similarity (67%) between the FadD19$_{DSM43269}$ and FadD19$_{H37Rv}$ enzymes further suggests that FadD19$_{H37Rv}$ has an *in vivo* role in sterol metabolism of *Mycobacterium tuberculosis* H37Rv.

INTRODUCTION

Phytosterols are among the most abundant sterols in nature and are mineralized from decaying plant material by soil bacteria, e.g. *Rhodococcus* strains belonging to the class of *Actinobacteria* (Watanabe *et al*., 1987). Microbial degradation of sterol molecules involves elimination of the alkyl side chain and opening of the polycyclic steroid nucleus (Fig. 1; Arima *et al*., 1969; Marsheck *et al*., 1972; Sih, 1962; Sih *et al*., 1968a, 1968b). The order of these two processes varies between bacterial genera, and even between members of the same genus (Rosłoniec *et al*., 2009).

Prior to their degradation, sterols are actively transported into the actinobacterial cells by the Mce4 steroid transporter (Mohn *et al*., 2008; Pandey and Sassetti, 2008; van der Geize *et al*., 2007). Following uptake by *R. jostii* RHA1 cells, the cytochrome P450 monooxygenase CYP125 initiates degradation of the sterol side chain by catalyzing hydroxylation of the C26 or C27 carbon (Fig. 1A and 1B, reaction 1; Capyk *et al*., 2009; McLean *et al*., 2009; Ouellet *et al*., 2010; Rosłoniec *et al*., 2009). After complete oxidation of the hydroxylated terminal carbon atom to its carboxylic acid intermediate (Fig. 1, compounds III and IX), the sterol side chain is shortened by a process similar to β-oxidation of fatty acids. This process is initiated by an ATP-dependent sterol/steroid-CoA ligase (Fig. 1A and 1B, reaction 3; Chen, 1985; Sih *et al*., 1968a, 1968b) catalyzing the coenzyme A (CoA) activation of the C26-carboxylic acid intermediates (Fig. 1, compounds III and IX). An ATP-dependent steroid-CoA ligase (65 kDa) from *Mycobacterium* sp. NRRL B3805 was purified to near homogeneity and shown to be highly specific towards C26-carboxylic sterols (Chen, 1985). The gene encoding this activity remained unknown, however. Thiolytic cleavage results in shortening of the cholesterol side chain in several rounds of β-oxidation via C24 (Fig. 1A, compound IV) and C22 (Fig. 1A, compound V) intermediates and concomitant release of propionyl-CoA and acetyl-CoA, respectively. This process has been elucidated at the biochemical level in strains of *Mycobacterium* and *Nocardia*, and proceeds differently for β-sitosterol compared to cholesterol (Fig. 1; Fujimoto *et al*., 1982a, 1982b; Marsheck *et al*., 1972; Martin and Wagner, 1976; Sih *et al*., 1968a, 1968b; Szentirmai, 1990).

Knowledge on genes involved in microbial degradation of the alkyl sterol side chain is extremely limited. Recently, a wealth of information on sterol catabolism was provided by the identification of a cholesterol catabolic gene cluster in *R. jostii* RHA1, encoding several enzymes with predicted functions in β-oxidation (van der Geize *et al*., 2007). A subset of these genes (i.e. *ro04683-ro04694*) is located proximal to *cyp125*, encoding steroid 26-hydroxylase involved in sterol side chain degradation (Capyk *et al*., 2009; McLean *et al*., 2009; Ouellet *et al*., 2010; Rosłoniec *et al*., 2009; van der Geize *et al*., 2007). Aim of our work is to decipher the roles of these genes in sterol metabolism, particularly those genes encoding steps of the β-oxidation cycle of sterol side chain degradation. The strain RHA1 cholesterol transcriptome identified two candidate genes, i.e. *fadD17* (*ro04691*) and *fadD19* (*ro04689*), encoding putative acyl-CoA ligases, that were highly upregulated during growth on cholesterol (Fig. 2A; van der Geize *et al*., 2007). Previous studies on FadD17 and FadD19 homologs from *M. tuberculosis* H37Rv showed that *in vitro* they act as CoA ligases capable

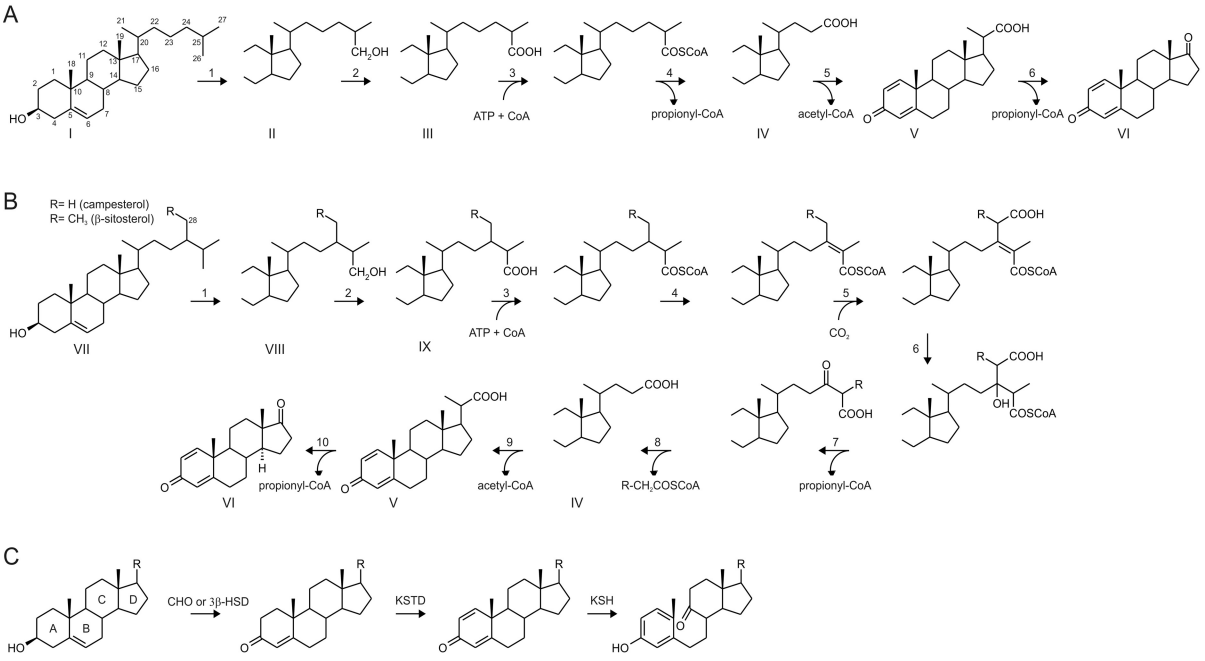


Fig. 1. Schematic overview of the side chain degradation pathways of (A) cholesterol, (B) β-sitosterol and campesterol, and (C) steroid ring opening in *Actinobacteria* (adapted from (Chen, 1985, Fujimoto *et al*., 1982b and van der Geize *et al*., 2007). Arrow numbering indicate reaction steps which are explained in the text. Abbreviations: CoA, coenzyme A, CHO, cholesterol oxidase, 3β-HSD, 3β-hydroxysteroid dehydrogenase, KSTD, 3-ketosteroid Δ1-dehydrogenase, KSH, 3-ketosteroid 9α-hydroxylase.

of activation of long-chain fatty acids (Trividi *et al*., 2004). Their *in vivo* physiological roles remained unknown. Only a few examples of sterol side chain degradation genes have been reported, including *baiB* of *Eubacterium* sp. strain VPI 12708 and *caiA* of *Pseudomonas* sp. Chol1, encoding a CoA ligase and an acyl-CoA dehydrogenase, respectively, involved in cholic acid degradation (Birkenmaier *et al*., 2007; Mallonee *et al*., 1992). Also, *fadA5* of *M. tuberculosis* H37Rv was shown to encode a thiolase with a role in cholesterol side chain degradation (Nesbitt *et al*., 2009).

Molecular characterization studies on sterol side chain degradation are generally hampered by the lack of a stable, genetically accessible strain with inactivated steroid ring degradation but capable of selective sterol side chain degradation. Interestingly, *Rhodococcus rhodochrous* IFO3338 (= DSM43269) is amenable to genetic manipulation and able to selectively degrade the sterol side chain in the presence of iron chelators, which inhibit 3-ketosteroid 9α-hydroxylase (KSH) activity (Arima *et al*., 1978; Rosłoniec *et al*., 2009). KSH is a two-component enzyme system, consisting of a terminal oxygenase KshA and a ferredoxin-reductase KshB, and a key-enzymatic step in steroid ring opening during microbial steroid catabolism (Petrusma *et al*., 2009; van der Geize *et al*., 2002a, 2008b). We genetically mimicked the chemical inhibition of KSH by iron chelators by constructing a stable multiple *kshA* gene deletion null mutant strain of *R. rhodochrous* DSM43269 (strain RG32) lacking KSH activity (Petrusma *et al*., in preparation). Mutant strain RG32 was completely blocked in

steroid ring degradation and capable of selective sterol side chain degradation, thereby accumulating 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC) and 1,4-androstadiene-3,17-dione (ADD) (Fig. 1, compounds V and VI) from sterols. The strain RG32 phenotype was used as a tool to assess the role of various genes in sterol side chain degradation. We identified *fadD19* as an essential gene for the degradation of the C24 branched-chain sterols campesterol and β-sitosterol, but not cholesterol. Direct enzyme activity measurement of heterologously produced FadD19 of DSM43269 conclusively identified its function as a steroid-CoA ligase in view of its ability to activate steroid carboxylic acid side chains.

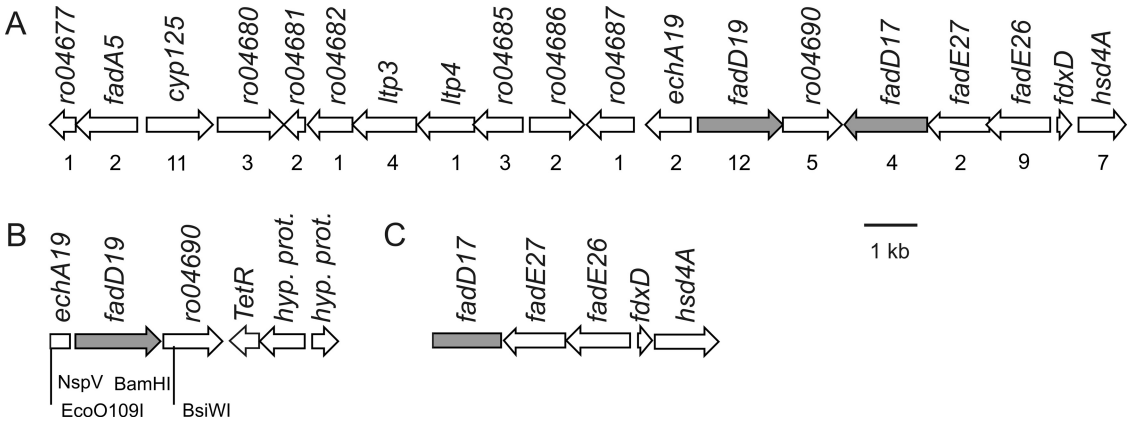


Fig. 2. (A) β-Oxidation gene cluster comprised of *ro04677-ro04695* within the cholesterol catabolic gene cluster of *R. jostii* RHA1 (McLeod *et al*., 2006;, van der Geize *et al*., 2007). The fold change of expression during growth on cholesterol as compared to pyruvate is also shown (van der Geize *et al*., 2007). Panel (B) and (C) show the genetic organization of homologs of *ro04688-ro04690* and *ro04691-ro04695* of *R. rhodochrous* DSM43269, respectively.

RESULTS

Deletion of the *ro04683-ro04694* gene subset of the cholesterol gene cluster in *Rhodococcus jostii* RHA1 does not affect growth on sterols

To assess the role of the subset of genes (*ro04683-ro04694*) within the cholesterol gene cluster predicted to be involved in sterol side chain degradation, mutant strain MW1 of strain RHA1 was constructed, carrying a deletion of in total 12 genes, i.e. *ro04683-ro04694* (Fig. 2). Growth experiments of mutant strain MW1 in mineral medium supplemented with 1 g l^{-1} cholesterol or β-sitosterol resulted in protein contents of 43 (± 4) and 53 (± 1) mg l^{-1}, respectively. These yields were comparable to those obtained with wild type RHA1 (56 (± 3) and 56 (± 4) mg l^{-1}, respectively).

Deletion of this set of genes thus did not block growth on cholesterol or β-sitosterol. We anticipated that sterol side chain degradation might be impaired in mutant MW1, but that growth on the sterol substrates could still be sustained through steroid ring oxidation, similar to what was previously observed for mutant *cyp125* of strain RHA1 (Rosłoniec *et al*., 2009). To specifically study the role of this set of genes, a strain is required blocked in ring

degradation but able to perform selective sterol side chain degradation. To block steroid ring degradation in strain RHA1, we attempted to inactivate all four *kshA* homologs (*ro02490*, *ro04538*, *ro05811*, *ro09003*) in strain RHA1 (van der Geize *et al*., 2007) by unmarked gene deletion mutagenesis (van der Geize *et al*., 2001). Following inactivation of strain RHA1 *ro05811* and *ro04538* we failed in subsequently deleting *ro09003*, for reasons unknown. The two-fold *kshA* deletion mutant of strain RHA1 was still able to convert sterols, did not accumulate steroid pathway intermediates, and thus did not perform selective side chain degradation (data not shown). These results prompted us to use another, more suitable *Rhodococcus* strain for studying the role of this set of genes in sterol side chain degradation, i.e. *R. rhodochrous* DSM43269.

Construction and characterization of *kshA* null mutant strain *R. rhodochrous* RG32 performing selective sterol side chain degradation

Wild type strain *R. rhodochrous* DSM43269 has five *kshA* homologs; the cloning of these homologs and the biochemical characterization of the encoded KshA proteins are described elsewhere (Petrusma *et al*., 2009; Petrusma *et al*., in preparation). The cloned DNA fragments of wild type *R. rhodochrous* DSM43269 were used to construct a five-fold *kshA* null mutant, designated strain RG32. Strain RG32 was blocked in growth on mineral medium containing 4-androstene-3,17-dione (AD) and 1,4-androstadiene-3,17-dione (ADD) (Petrusma *et al*., in preparation). Whole cell biotransformations of AD by strain RG32 resulted in 69% molar conversion into ADD (Table 1, Fig. 1C: KSTD activity), which was not accumulated in detectable amounts by wild-type strain DSM43269 (data not shown), confirming impaired KSH activity (van der Geize *et al*., 2002a). Next, mutant strain RG32 was tested in whole cell biotransformations for its ability to perform selective side chain degradation on a range of sterols/steroids.

All compounds tested were converted into ADD (Fig. 1, compound VI) and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC; Fig. 1, compound V) in molar ratios of 1-7% and 50-80%, respectively, depending on the substrate used (Table 1). Thus, mutant strain RG32

Table 1. Whole cell bioconversion of sterols/steroids into ADD and 1,4-BNC by *R. rhodochrous* mutant strain RG32 after 72 h of incubation (NA = not applicable, - = not detected). The data represent the molar % conversion as averages of triplicate experiments; standard deviations are shown between brackets.

Steroid substrate	ADD (%)	1,4-BNC (%)
4-androstene-3,17-dione	69 (± 9)	NA
23,24-bisnor-5-cholene-22-oic acid-3β-ol	-	60 (± 21)
campesterol	1 (± 0.5)	49 (± 9)
5-cholenic acid-3β-ol	1 (± 0.3)	77 (± 4)
4-cholestene-3-one	4 (± 1)	79 (± 2)
1,4-cholestadiene-3-one	3 (± 1)	71 (± 4)
cholesterol	3 (± 1)	73 (± 12)
β-sitosterol	7 (± 2)	67 (± 7)

is completely blocked in steroid ring degradation and capable of selective sterol side chain degradation, enabling analysis of the putative sterol side chain degradation genes.

Cloning homologous cholesterol catabolic genes from *R. rhodochrous* DSM43269

Sequence data on the cholesterol catabolic gene cluster of *R. rhodochrous* DSM43269, necessary to perform mutational analysis of genes involved in sterol side chain degradation in strain RG32, was not available yet. Therefore, homologs of *R. jostii* RHA1 genes predicted to be involved in sterol side chain degradation were cloned from *R. rhodochrous* DSM43269. First, a genomic library of strain DSM43269 (Petrusma *et al*., 2009) was screened using degenerate PCR primers based on amino acid sequences that were highly conserved among actinobacterial homologs of strain RHA1 Ro04690 and FadE26 (Ro04693) (Fig. 2A). Library screening by PCR resulted in the isolation of two separate clones that were sequenced and analyzed (Fig. 2B and 2C). A 5.4 kb genomic fragment (Fig. 2B) carried $ro04690_{DSM43269}$, encoding a protein displaying 85% amino acid sequence identity with $Ro04690_{RHA1}$. Two genes located immediately upstream of $ro04690_{DSM43269}$ encode proteins sharing 74% and 76% amino acid sequence identity with the deduced amino acid sequences of *ro04688* (partial sequence) and *fadD19* in strain RHA1, respectively. Three genes located downstream of $ro04690_{DSM43269}$ encode proteins showing the highest similarities to Ro03510 (23%), Ro04422 (25%) and Ro01580 (56%) of strain RHA1, respectively. Bioinformatic analysis of the sequence of a 5.5 kb insert (Fig. 2C) confirmed the presence of $fadE26_{DSM43269}$, whose gene product displayed 86% amino acid sequence identity to its counterpart in strain RHA1. Moreover, the fragment contained genes that encode proteins with 46-70% amino acid sequence identities to *fadD17* (partial sequence), *fadE27*, *ro04694* and *hsd4A* of strain RHA1, exhibiting identical genetic organizations (Fig. 2). The two cloned genomic fragments of strain DSM43269 did not show sequence overlap.

FadD19 is essential for C24-branched chain sterol side chain degradation

To substantiate a role of $fadD17_{DSM43269}$, $fadD19_{DSM43269}$, $fadE26_{DSM43269}$, $fadE27_{DSM43269}$ and $ro04690_{DSM43269}$ in sterol side chain degradation, gene inactivation mutants of strain RG32 were constructed. Biotransformations of cholesterol with whole-cell cultures of these mutants revealed that all mutants displayed parental phenotypes, with similar amounts of ADD and 1,4-BNC accumulating from cholesterol as observed with strain RG32 (Table 2). Conceivably, the putative CoA ligases *fadD17* and *fadD19* are functional homologs able to cross complement each others during cholesterol catabolism. To test this hypothesis, a Ω*fadD17* inactivation mutant was constructed in strain RG32Δ*fadD19* and its ability to degrade cholesterol was tested by whole cell biotransformation. Cholesterol degradation also was unaffected in RG32Δ*fadD19* Ω*fadD17*, displaying conversion rates similar to that of strain RG32 (Table 2).

FadE26, FadE27 and Ro04690 all three belong to the superfamily of acyl-CoA dehydrogenases (Daubner *et al*., 2002). We therefore also constructed a double gene

inactivated mutant strain RG32Δ*ro04690*$_{DSM43269}$ Δ*fadE26* and a triple mutant RG32Δ*ro04690*$_{DSM43269}$ Δ*fadE26 ΩfadE27* to prevent possible cross-complementation. However, all these mutants were unaffected in cholesterol side chain degradation and displayed parent strain RG32 phenotypes (Table 2). Although clearly upregulated in strain RHA1 cells grown on cholesterol (Fig. 2A), these *R. rhodochrous* DSM42369 genes either are not involved in cholesterol side chain degradation, or the reaction can be performed by additional isoenzymes encoded by strain DSM42369.

Table 2. Whole-cell bioconversion of sterols by *R. rhodochrous* RG32 and mutants thereof into ADD and 1,4-BNC, measured after 72 h of incubation (ND = not determined, - = not detected). The data represent the molar % conversion averages of duplicate experiments; standard deviations are shown between brackets.

	cholesterol		β-sitosterol		campesterol	
strain	ADD (%)	1,4-BNC (%)	ADD (%)	1,4-BNC (%)	ADD (%)	1,4-BNC (%)
RG32	3 (±1)	73 (±12)	7 (±2)	67 (±7)	1 (±0)	49 (±9)
RG32Δ*ro04690*$_{DSM43269}$	4 (±1)	74 (±8)	7 (±4)	54 (±9)	ND	ND
RG32Δ*ro04690*$_{DSM43269}$ Δ*fadE26*	3 (±1)	72 (±4)	6 (±4)	61 (±9)	ND	ND
G32Δ*ro04690*$_{DSM43269}$ Δ*fadE26 ΩfadE27*	3 (±2)	70 (±6)	7 (±2)	65 (±6)	ND	ND
RG32Ω*fadD17*	3 (±2)	74 (±9)	8 (±2)	64 (±5)	ND	ND
RG32Δ*fadD19*	6 (±2)	64 (±4)	-	-	-	-
RG32Δ*fadD19* +pCOMP*fadD19*$_{DSM43269}$	ND	ND	6 (±1)	60 (±4)	ND	ND
RG32Δ*fadD19 ΩfadD17*	6 (±2)	69 (±1)	-	-	ND	ND

Next, all mutants were tested in whole cell biotransformations with β-sitosterol, a C24-ethyl branched phytosterol (Fig. 1B). Interestingly, strain RG32Δ*fadD19* was blocked in its ability to convert the side chain of β-sitosterol (Table 2). All other mutants displayed parental strain phenotypes, accumulating similar amounts of ADD and 1,4-BNC from β-sitosterol as strain RG32 (Table 2). Further analysis of the RG32Δ*fadD19* mutant revealed that also the degradation of campesterol, a C24-methyl branched sterol, was blocked (Table 2).

When incubated with β-sitosterol, strain RG32Δ*fadD19* accumulated metabolites that were not detected in strain RG32 (Fig. 3). These metabolites corresponded to ring oxidized derivatives of the β-sitosterol mixture, i.e. 4-sitostene-3-one and 4-campestene-3-one, as was confirmed by reference samples. Mutant strain RG32Δ*fadD19* regained parental phenotype when complemented with *fadD19*$_{DSM43269}$ under control of the *aphII* promoter (Table 2). This excluded the possibility that β-sitosterol side chain degradation in RG32Δ*fadD19* was blocked by polar effects rather than by inactivation of *fadD19* directly.

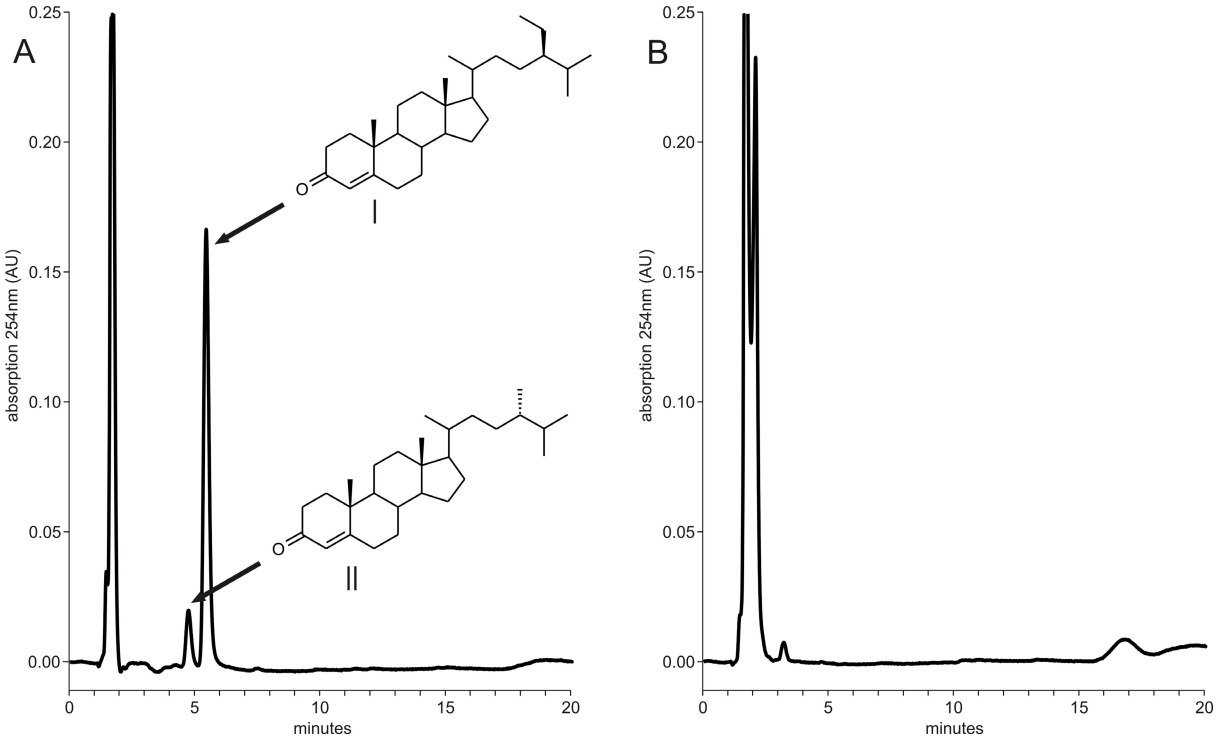


Fig. 3. HPLC graphs of steroid extracts from β-sitosterol bioconversions after three days of incubation. (A) Profile of mutant strain RG32Δ*fadD19* showing the accumulation of metabolites (degradation pathway intermediates) that based on identical HPLC retention times were identified as 4-sitostene-3-one (I) and 4-campestene-3-one (II), and (B) profile of parent strain RG32.

FadD19$_{DSM43269}$ displays sterol CoA-ligase activity towards C26-oic acid steroids

The gene inactivation studies in strain RG32 showed that *fadD19* is essential for side chain degradation of branched chain sterols. To substantiate that *fadD19* encodes a steroid-CoA ligase, FadD19$_{DSM43269}$ activity was tested. Heterologous expression of the protein was achieved using *E. coli* strain BL21(DE3) carrying pET15b*fadD19$_{DSM43269}$*. Cell-free extracts (CFE) derived from *E. coli* cultures expressing FadD19$_{DSM43269}$ were active towards the side chain of 5-cholestene-26-oic acid-3β-ol, using ATP and CoA as co-substrates (Fig 4A). A negative control, consisting of the same components, but with CFE obtained from cultures of the same *E. coli* strain carrying empty pET15b vector, did not show product formation (Fig. 4B). Incubations of CFE containing FadD19$_{DSM43269}$ also resulted in the formation of reaction products with 3-oxo-4-cholestene-26-oic acid or 5-cholenic acid-3β-ol in a steroid substrate, ATP and CoA dependent manner (Fig. 4A). Omitting any of these reaction components resulted in loss of enzyme activity, indicating that these are all essential components. Therefore, the products formed are most likely the CoA esters of 5-cholestene-26-oic acid-3β-ol and 3-oxo-4-cholestene-26-oic acid, confirming that FadD19$_{DSM43269}$ functions as a steroid-CoA ligase. Activity towards C24-branched chain steroids with a terminal carboxylic acid could not be tested, since these compounds were not (commercially) available.

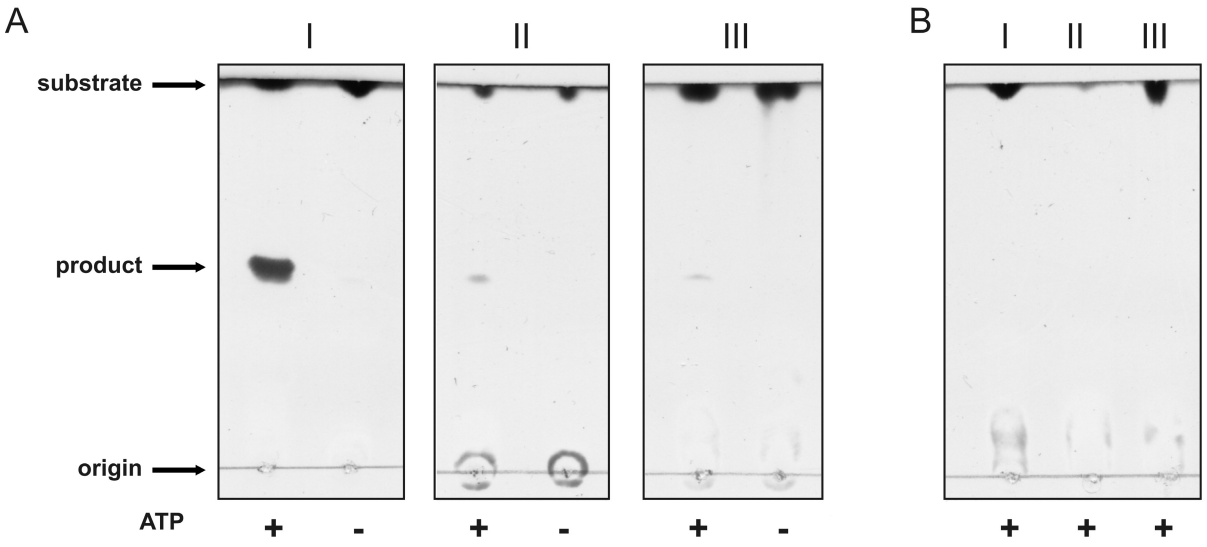


Fig. 4. (A) TLC analysis of the reactions of cell-free extracts (CFE) of *E. coli* BL21(DE3) cells expressing *fadD19*$_{DSM43269}$, incubated with: (I) 5-cholestene-26-oic acid-3β-ol, (II) 3-oxo-4-cholestene-26-oic acid and (III) 5-cholenic acid-3β-ol. (B) TLC analysis of reactions of CFE of BL21 cells containing empty plasmid pET15b, using the same sterol substrates (I-III). All incubations were performed for 4 hours at 30°C and contained the co-substrates CoA and Mg^{2+}. ATP was either included (+) or omitted (-) as a negative control. Product formation was also dependent on the presence of CoA (data not shown).

Bioinformatic analysis and phylogeny of *R. jostii* RHA1 CoA ligases

Database searches served to identify homologs of both FadD19$_{RHA1}$ and FadD17$_{RHA1}$ in closely related sterol degrading *Actinobacteria*, including *Nocardia farcinica* IFM 10152 (Nfa5290 and Nfa24170: with 75 and 47% identity, respectively), *Mycobacterium smegmatis* str. MC2 155 (MSMEG_5914 and MSMEG_5908: with 67 and 57% identity, respectively), *M. tuberculosis* H37Rv (Rv3515c and Rv3506: with 67 and 56% identity, respectively) and *Rhodococcus equi* 103S and ATCC33707 (HMPREF0724_4865 and HMPREF0724_4863: with 77 and 72% identity, respectively). A comparison of FadD19$_{RHA1}$ and FadD17$_{RHA1}$ with the amino acid sequence of the cholic acid-CoA ligase, encoded by *baiB*, revealed relatively low similarities (24 and 20% identity, respectively). All sequences in the strain RHA1 genome annotated as CoA ligase or CoA synthetase were compared and depicted in a phylogenetic tree (Fig. 5). This revealed that the strain RHA1 genome contains closely related homologs of FadD17$_{RHA1}$ and FadD19$_{RHA1}$, i.e. Ro05822 (44% identity) and Ro04675 (49% identity), respectively. These putative enzymes thus may have a similar steroid-CoA ligase activity, resulting in a metabolic redundancy for activation of the side chain of 5-cholestene-C26-oic acid-3β-ol metabolites. Conversely, redundancy for C24-branched sterol side chain activation is absent.

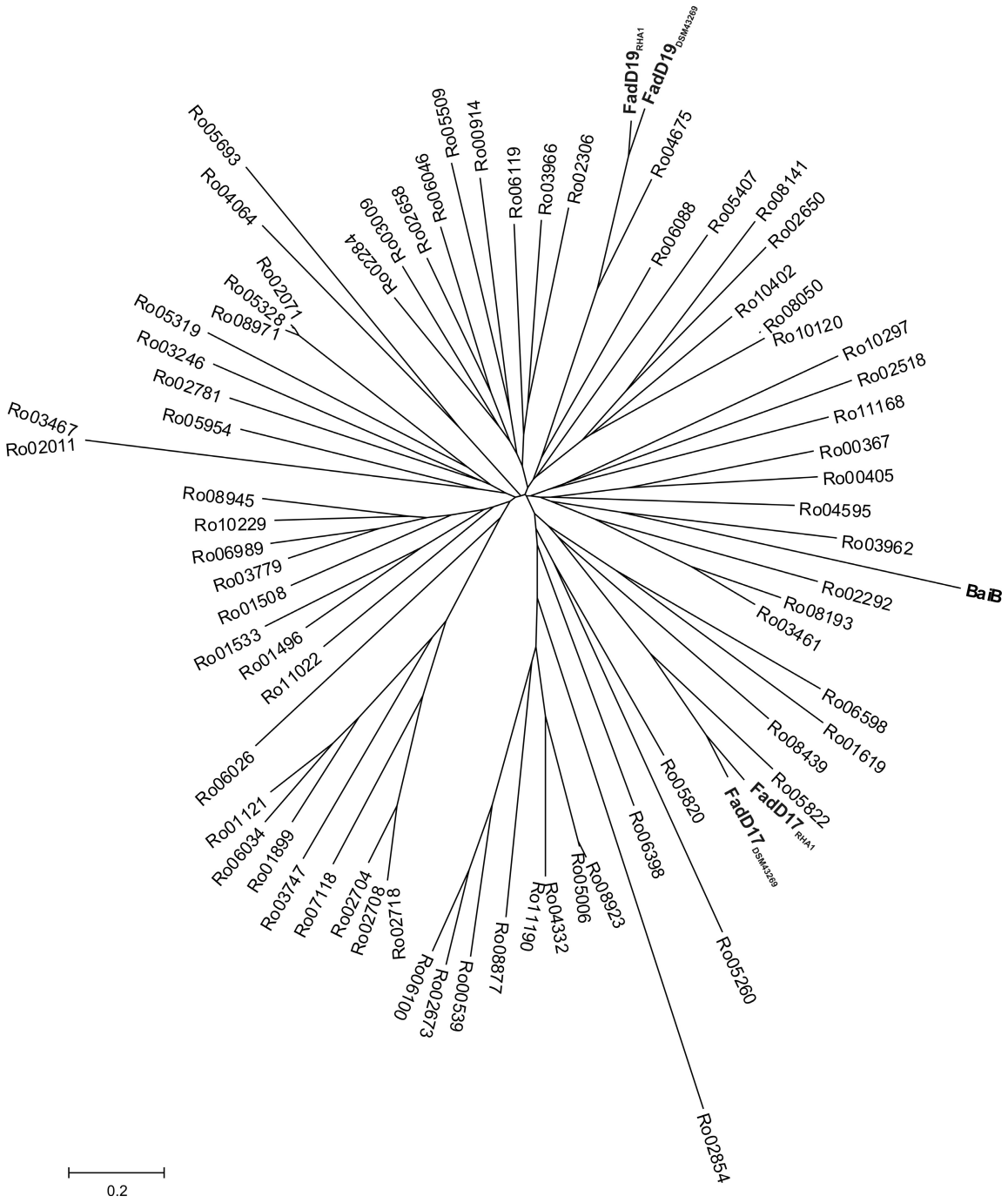


Fig. 5. Phylogenetic tree of 73 amino acid sequences that are annotated as (AMP-forming) CoA ligases or CoA synthetases in the *R. jostii* RHA1 genome (McLeod *et al*., 2006), including FadD17 and FadD19 from both strain RHA1 and *R. rhodochrous* DSM43269, and the cholic acid-CoA ligase BaiB from *Eubacterium* sp. strain VPI 12708 (Mallonee *et al*., 1992).

DISCUSSION

Knowledge on genes and enzymes involved in microbial degradation of the alkyl sterol side chain has remained extremely limited. Our research focuses on the characterization of genes encoding the initial steps in microbial sterol side chain degradation. Recently, we reported the identification and characterization of CYP125 of *R. jostii* RHA1 and *R. rhodochrous* DSM43269 as a steroid 26-monooxygenase essential for sterol side chain

degradation (Rosłoniec *et al*., 2009). The present study identified FadD19 *of R. rhodochrous* DSM43269 as a steroid-CoA ligase, essential for the degradation of C24-branched chain sterols. An unmarked *fadD19* deletion mutant in *kshA* null mutant strain RG32 (strain RG32Δ*fadD19*) was blocked in side chain degradation of the C24-branched sterols β-sitosterol and campesterol, since no detectable amounts of 1,4-BNC or ADD were detected (Table 2). In biotransformations of β-sitosterol with whole cells of mutant RG32Δ*fadD19*, but not with cells of parent strain RG32, steroid metabolites accumulated whose HPLC elution profiles corresponded to derivatives of β-sitosterol and campesterol with oxidized steroid rings These results therefore suggest that in strain RG32Δ*fadD19* no sterol side chain degradation occurs, but only steroid ring oxidation. From these experiments we conclude that $FadD19_{DSM43269}$ catalyzes an essential step in the side chain degradation of C24-branched sterols in *R. rhodochrous* DSM43269. Surprisingly, degradation of the side chain of cholesterol was not affected in strain RG32Δ*fadD19* and accumulation of 1,4-BNC and ADD was comparable to that in parent strain RG32 (Table 2). In *R. jostii* RHA1, *fadD17* is located proximal to *fadD19* within the cholesterol catabolic gene cluster and highly up-regulated during growth on cholesterol. Thus, *fadD17* is a likely candidate to encode a steroid-CoA ligase involved in cholesterol side chain activation (van der Geize *et al*., 2007). Mutant strains with an inactivated $fadD17_{DSM43269}$ gene in both RG32 and RG32Δ*fadD19,* however, were still capable of transformation of the cholesterol side chain. Therefore, it remains elusive whether side chain activation during cholesterol catabolism is performed by multiple redundant CoA ligases, or that a yet unidentified, highly specific enzyme catalyzes this reaction *in vivo*.

Direct enzyme activity measurements, using *E. coli* cell-free extracts expressing $FadD19_{DSM43269}$ showed that the enzyme indeed was able to activate the side chains of 5-cholestene-26-oic acid-3β-ol, 3-oxo-4-cholestene-26-oic acid and 5-cholenic acid-3β-ol, in an ATP and CoA dependent manner (Fig. 4). Since cell-free extracts obtained from the same *E. coli* strain but carrying an empty plasmid did not result in detectable amounts of product formation with these substrates (Fig. 4B), product formation as observed in Fig. 4A must originate from $FadD19_{DSM43269}$ activity. The intensities of the reaction product spots suggest that $FadD19_{DSM43269}$ has substrate preference for steroids possessing longer side chains over shorter ones. Furthermore, substrates with a 3-hydroxy-5-ene ring structure may be better substrates than their oxidized forms. However, it is difficult to quantitatively compare the reaction products due to the discrepancy in spot intensity between the 3-hydroxy-5-ene and 3-oxo-4-ene substrates at the same substrate concentration. Additional biochemical studies are thus required to assess the substrate specificity of $FadD19_{DSM43269}$.

Despite the fact that $FadD19_{DSM43269}$ was able to activate 5-cholestene-C26-oic acid-3β-ol and derivatives thereof *in vitro*, deletion of $fadD19_{DSM43269}$ in strain RG32 did not affect cholesterol degradation. Therefore, it is most plausible that, besides *fadD19* and possibly *fadD17*, other isoenzymes are encoded by the DSM43269 genome that can complement 5-cholestene-C26-oic acid-3β-ol activation in this mutant. The genome of *R. jostii* RHA1 (McLeod *et al*., 2006) contains at least 73 genes annotated to encode putative CoA ligases or

CoA synthetases. Phylogenetic analysis of these (Fig. 5) showed that in strain RHA1, both $FadD17_{RHA1}$ and $FadD19_{RHA1}$ have relatively close homologs, *i.e.* Ro05822 and Ro04675, respectively, displaying 44 and 49% amino acid sequence identities, respectively. In all other available rhodococcal genomes homologs of both Ro05822 and Ro04675 are present, which are all reciprocal best hits. Therefore, it is plausible that these homologs are also encoded by the strain DSM43269 genome and account for the lack of phenotype found with our mutants in whole cell biotransformations with cholesterol. CoA ligases are known to be promiscuous enzymes, that often have overlapping substrate specificities (Arora *et al*., 2005, Pei *et al*., 2003), further supporting the hypothesis that the genomes of strains RHA1 and DSM43269 may encode multiple enzymes with the ability to activate the side chain of 5-cholestene-26-oic acid-3β-ol or derivatives thereof (e.g. its 3-oxo-4-ene metabolite). Consistent with this hypothesis, inactivation of multiple genes encoding putative acyl-CoA dehydrogenases (i.e. $fadE26_{DSM43269}$, $fadE27_{DSM43269}$ and $ro04690_{DSM43269}$) in strain RG32 did not hamper cholesterol nor β-sitosterol side chain degradation (Table 2). Rhodococcal genomes are very rich in genes encoding β-oxidation enzymes, making it very likely that enzymatic redundancy for these functions exist. However, the possibility that the $fadE26_{DSM43269}$, $fadE27_{DSM43269}$ and $ro04690_{DSM43269}$ genes in fact do not have a role in sterol side chain degradation cannot be ruled out at the moment.

Interestingly, there is apparently no redundancy for activation of branched sterol side chains, since deletion of $fadD19_{DSM4269}$ alone in strain RG32 resulted in impaired degradation of branched chain sterols. $FadD19_{DSM43269}$ may be the only enzyme capable to activate the side chains of branched phytosterols, which possess a bulkier side chain than cholesterol. However, it cannot be excluded that a homolog (e.g. Ro04675) is encoded by the strain DSM43269, capable to catalyze the same reaction, which is not expressed in mutant RG32Δ*fadD19* under the growth conditions used.

Mutant strain *R. jostii* MW1, devoid of the entire subset of predicted sterol side chain degradation genes including *fadD19* (i.e. *ro04683-ro04694*), retained the ability to grow on sterols. However, a blocked growth on β-sitosterol of MW1 would have been expected, in light of the inability of RG32Δ*fadD19* to transform β-sitosterol in bioconversions. Growth of mutant MW1 on sterols most probably is still supported by ongoing ring oxidation following the initiation of side chain degradation mediated by $CYP125_{RHA1}$ (Rosłoniec *et al*., 2009). We have previously shown that initiation of sterol side chain degradation in strain RHA1 by CYP125 precedes sterol ring oxidation (Rosłoniec *et al*., 2009). Mutant strain RHA1Δ*cyp125* is blocked in initiation of sterol side chain degradation, but retained the ability to grow on 3-keto sterols through steroid ring oxidation (Rosłoniec *et al*., 2009). Since mutant strain MW1 still contains *cyp125*, initiation of side chain degradation is not blocked and complete ring oxidation, followed by degradation of the complete molecule, may account for the ability of strain MW1 to grow normally on 3-hydroxy sterols, despite the fact that the side chain degradation genes are lacking.

The results described in our current study are consistent with a role of FadD19 of *R. rhodochrous* DSM43269 as a steroid-CoA ligase reaction with an *in vivo* function in the

degradation of C24-branched sterol side chains. $FadD19_{H37Rv}$ of the human pathogen *M. tuberculosis* H37Rv, encoded by *rv3515c*, possesses CoA ligase activity and catalyzes the formation of acyl-CoA from long-chain fatty acids (Trivedi *et al*., 2004). Blast searches with $FadD19_{DSM43269}$ and $FadD19_{RHA1}$ against the H37Rv genome showed that they both have the highest sequence identity with $FadD19_{H37Rv}$; $FadD19_{RHA1}$ and $FadD19_{H37Rv}$ also were best reciprocal hits. The high similarity (67%) between these enzymes suggests an *in vivo* role for $FadD19_{H37Rv}$ in sterol metabolism of *M. tuberculosis* H37Rv.

This is the first report on the identification and characterization of a CoA ligase with an *in vivo* role in sterol side chain degradation. The study also demonstrated the successful use of mutant strain *R. rhodochrous* RG32, blocked in sterol ring degradation, as a tool to study the process of microbial sterol side chain degradation.

MATERIALS AND METHODS

Bacterial strains, growth conditions, plasmids and chemicals

Bacterial strains and plasmids used in this study are listed in supplemental Table S1. *E. coli* strains were grown in Luria-Bertani (Sigma, Zwijndrecht, Netherlands) medium at 37°C. Liquid cultures were grown with shaking (220 rpm). Ampicillin (100 μg ml^{-1}), kanamycin (25 μg ml^{-1}) or apramycin (50 μg ml^{-1}) were added to the medium if appropriate after autoclaving. For growth on solid medium, Difco™ Agar (Becton, LePont de Claix, France) was added at a final concentration of 1.5% (w/v). *Rhodococcus* strains were grown at 30°C with constant agitation (220 rpm) for liquid cultures. LBP medium (1% NaCl, 0.5% yeast extract (BBL™; Becton, LePont de Claix, France) and 1% Bacto™ Peptone (Becton, LePont de Claix, France)) was used for intergenic DNA conjugation (van der Geize *et al*., 2001). Kanamycin was added to final concentrations of 200 μg ml^{-1} and 25 μg ml^{-1} for *R. jostii* RHA1 and *R. rhodochrous* DSM43269, respectively. Sucrose sensitivity was tested on LBP agar supplemented with sucrose (10% (w/v)). *R. rhodochrous* DSM43269 and mutant strains derived were grown in LB medium; when appropriate, kanamycin (25 μg ml^{-1}) or apramycin (50 μg ml^{-1}) was added.

The ability of *R. jostii* RHA1 and mutant MW1 to grow on sterols was tested in mineral medium (MM) (Masai *et al*., 1995), supplemented with 2.5 mM of either cholesterol or β-sitosterol inoculated (1:100) by three day old pre-cultures grown MM supplemented with pyruvate (20 mM). Before autoclaving, sterols were dispersed by sonication for 10 min using a Branson Ultrasonic sonicator at an output of 10 μm. Biomass production was determined by measuring total protein content, as described (Rosłoniec *et al*., 2009). Cholesterol, campesterol and coenzyme A trilithium salt were obtained from Sigma-Aldrich. β-Sitosterol was obtained from Acros Organics. 5-Cholestene-26-oic acid-3β-ol was synthesized at Schering-Plough, Oss (Rosłoniec *et al*., 2009). 4-Androstene-3,17-dione, 1,4-androstadiene-3,17-dione, 23,24-bisnor-5-cholene-22-oic acid-3β-ol, 3-oxo-23,24-bisnor-4-cholene-22-oic acid, 4-cholestene-3-one, 1,4-cholestadiene-3-one, 5-cholenic acid-3β-ol and 3-oxo-4-cholestene-26-oic acid were obtained from Steraloids. ATP was purchased at Duchefa, Haarlem, The Netherlands.

General molecular techniques

Recombinant DNA techniques were performed according to standard protocols (Sambrook *et al*., 1989). DNA modifying enzymes (restriction enzymes ligases and DNA polymerases) were purchased from Roche (Mannheim, Germany), New England Biolabs (Beverly, Mass. USA) or Fermentas (St. Leon-Rot, Germany) and were used according to the manufacturer's protocols. All PCR products were obtained using *Pwo* polymerase (Roche) using standard conditions: 30 cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 1 min. DNA fragments were purified from agarose gel using Sigma Genelute™ gel extraction kit.

Screening of a R. rhodochrous DSM43269 genomic library for ro04690 and fadE26 homologs

Degenerate *ro04690*^{DEG} primers (Table S2) were based on conserved amino acid sequences HTGE(I/V)A(S/T)M (aa 193-200) and PLPMPLQ (aa 306-312) in Ro04690. Degenerate *fadE26*^{DEG} primers (Table S2) were based on conserved amino acid sequences V(I/V)NGQKMW (aa 153-160) and FGGGTNE(V/I) (aa 369-376) in FadE26$_{RHA1}$. A genomic library of *R. rhodochrous* DSM43269 (Petrusma *et al*., 2009) was screened by PCR using the *ro04690*^{DEG} and *fadE26*^{DEG} primers for clones containing the full length *ro04690* and *fadE26* genes, respectively.

Targeted disruption and unmarked gene deletions in Rhodococcus strains

Disruption and unmarked gene deletion mutants of *R. jostii* RHA1 and *R. rhodochrous* RG32 were constructed using the *sacB* counter-selection system (van der Geize *et al*., 2001).

Mutagenic plasmids for targeted gene deletion were constructed by PCR amplification of up- and downstream regions of the target genes using the primers listed in supplemental Table S2. The obtained amplicons were separately subcloned into *Eco*RV digested pBlueScript(II)KS. The amplicons were then ligated using a native restriction site from pBlueScript(II)KS (*Eco*RI or *Xba*I), and the restriction site that was introduced by PCR, being either *Bgl*II or *Nde*I. The genomic fragment was then isolated by digestion with either *Eco*RI or *Xba*I and *Hind*III and cloned into pK18*mobsac*B.

A construct for deletion of *fadD19*$_{DSM43269}$ in strain RG32 was made as follows: pRESQ-4690 was digested with *Eco*O119I and *Pst*I giving a 3.0 kb fragment, whose 5' overhang was filled in with Klenow and 3' overhang was removed by T4 DNA polymerase. Plasmid pK18*mobsac*B was digested with *Bam*HI and made blunt ended by Klenow fill-in and used for ligation with the blunt ended 3.0 kb fragment, giving pK18*fadD19*. A 751 bp deletion was introduced by digesting pK18*fadD19* with *Nsp*V and *Bam*HI, followed by 5' overhang fill-in using Klenow and self ligation, yielding mutagenic plasmid pDEL*fadD19*$_{DSM43269}$.

Mutagenic plasmids for targeted gene disruptions were obtained by PCR amplification of internal fragments of the target genes, using the primers listed in supplemental Table S2. The amplicons were subsequently ligated into *Sma*I digested pK18*mobsacB*.

Mutagenic plasmids were transferred to *E. coli* S17-1 by transformation and subsequently mobilized to *Rhodococcus* spp. by conjugational transfer as described (34). All mutants were

verified by PCR using specific primers (supplemental Table S2) to confirm deletion or disruption of the target gene(s).

Functional complementation of RG32ΔfadD19

The *fadD19*$_{DSM43269}$ gene was amplified by PCR using the PCR primers compfadD19-F and compfadD19-R. The obtained PCR product of 1,673 bp was digested with *Acc*65I and ligated into *Eco*RV/*Acc*65I digested pBs-Pkan, yielding plasmid pBs-Pkan-*fadD19*$_{DSM43269}$. The pBs-Pkan-*fadD19*$_{DSM43269}$ plasmid was digested with *Spe*I/*Acc65*I to obtain the 2.1 kb Pkan-*fadD19*$_{DSM43269}$ cassette that was subsequently ligated into *Spe*I/*Acc65*I digested pRESQ, resulting in plasmid pCOMP*fadD19*$_{DSM43269}$ that was used to transform electrocompetent cells of RG32Δ*fadD19* as described (Fernandes *et al*., 2001).

Whole-cell biotransformation of R. rhodochrous RG32 derived mutants and steroid analyses

Cell cultures of *R. rhodochrous* RG32 and derived mutant strains were grown overnight in liquid LB medium at 30°C with shaking (220 rpm) until mid-exponential phase. Stock solutions of sterols were prepared by dissolving in acetone at a concentration of 25 mM, and added to cultures at a final concentration of 0.5 mM. After further incubation for three days under the same conditions, samples were taken for HPLC analysis.

To analyze sterol conversions, 1 ml culture samples were taken and centrifuged (1 min 16,000 g) and 0.5 ml of supernatant was mixed with 2 ml methanol and filtered prior to analysis by HPLC. HPLC was performed on an Alltima C18 column (250 x 4.6 mm; Alltech, Deerfield, USA, 35°C) using a mobile phase consisting of methanol:water:formic acid (80:19:1) at a flow rate of 1 ml min^{-1} with UV detection at 254 nm. Authentic 1,4-androstadiene-3,17-dione and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid, obtained by incubating authentic 3-oxo-23,24-bisnor-4-cholene-22-oic acid with purified 3-ketosteroid dehydrogenase enzyme (Δ1-KSTD1; Knol *et al*., 2008), were used as references to enable quantification of the conversion rates. Steroids were extracted from culture broth using two volumes of ethylacetate. Samples were directly applied onto the same HPLC and detection system as described above, except that the mobile phase consisted of acetonitrile:tetrahydrofuran (75:25) and the flow rate was set at 2 ml min^{-1}. Authentic β-sitosterol (a mixture of β-sitosterol and campesterol) was incubated with cholesterol oxidase (Sigma) to obtain 4-sitostene-3-one and 4-campestene-3-one and used as reference sample.

Production of CoA ligase FadD19$_{DSM43269}$

The *fadD19*$_{DSM43269}$ gene was amplified by PCR from genomic DNA of *R. rhodochrous* DSM43269 using primer pair fadD19exp-F and fadD19Exp-R (supplemental Table S2). After subcloning of the amplicon in pBlueScript(II)KS, the gene was cloned into pET15b using *Nde*I and *Bam*HI, yielding pET15b*fadD19*$_{DSM43269}$ and introduced into competent *E. coli* BL21(DE3) cells for expression. Overexpression of FadD19$_{DSM43269}$ was obtained by growing 50 ml cultures in LB medium, supplemented with 0.5 M sorbitol, 100 mM ampicillin and 0.1 mM

IPTG for 24 h at 25°C with shaking (220 rpm). Cells were harvested by centrifugation (4,600 g, 15 min, 4°C). Cell pellets were resuspended in 50 mM Tris buffer pH 8.0 and cells were disrupted by sonication (10 × 30 sec at 10 micron with 30 sec intervals) on ice. To obtain cell-free extracts, cell debris were removed by centrifugation (40,000 g, 20 min, 4°C). Proteins were analyzed by SDS-PAGE analysis as described (25) and quantified using the Bradford method, using the Bio-Rad reagent and bovine serum albumin (BSA) as a standard.

Enzymatic activity of FadD19$_{DSM43269}$

Activity tests of FadD19$_{DSM43269}$ on various carboxylic acid sterols was analyzed by thin layer chromatography (TLC). Reaction mixtures consisted of 20 mM Tris pH 8.0, 5 mM $MgCl_2$, 0.5 mM of substrate and 2 mM ATP and/or 1 mM CoA, where applicable, in a final volume of 10 µl. Reactions were initiated by the addition of 1 µl of CFE containing FadD19$_{DSM43269}$ (~ 4 µg of total protein). Care was taken to use freshly produced enzyme from the same production batch in all experiments. CFE of *E. coli* BL21(DE3) carrying empty pET15b was used as negative control. After 4 hours of incubation at 30ºC, total reaction volumes were directly applied to F_{254} silica gel TLC plates (Merck, Darmstadt, Germany) and separated using a mobile phase consisting of 1-butanol:acetic acid:water (80:25:40) for 30 min. Spots were visualized using Cer-reagent consisting of 1% (w/v) $Ce(SO_4)_2$, 2.5% (w/v) $H_3[P(Mo_3O_{10})_4].H_2O$ and 8% (v/v) concentrated H_2SO_4 and developed by heating with a hot air gun.

Phylogenetic tree construction

Protein sequences of CoA ligases from *R. jostii* RHA1 were aligned using ClustalW; MEGA4.1 was used for phylogenetic tree construction (Tamura *et al*., 2007). Amino acid sequences, annotated as acyl-CoA ligases or acyl-CoA synthetases, were obtained from the RHA1 genome website (www.rhodococcus.ca) and used for phylogenetic tree construction. Furthermore, BaiB of *Eubacterium* sp. strain VPI 12708 (accession number P19409) was included in the phylogenetic tree.

DDBJ/EMBL/GENBANK DATABASE ACCESSION NUMBERS

The *R. rhodochrous* DSM43269 sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession numbers HM588719 and HM588720.

ACKNOWLEDGEMENTS

This work was financially supported by the IBOS Program (Integration of Biosynthesis & Organic Synthesis) of Advanced Chemical Technologies for Sustainability (ACTS).

REFERENCES

References are listed on pages 119-140.

CHAPTER 5

The *ltp3* and *ltp4* genes of *Rhodococcus rhodochrous* DSM43269 are essential for side chain degradation of C24-branched sterols

Maarten H. Wilbrink, Robert van der Geize and Lubbert Dijkhuizen

In preparation for submission

ABSTRACT

The *ltp3* and *ltp4* genes, annotated as putative 3-ketoacyl-CoA thiolases, enzymes with a role in β-oxidation, are located within a cholesterol catabolic gene cluster in *Rhodococcus jostii* RHA1. Deletion of a large set of genes, including *ltp3* and *ltp4*, in strain RHA1 was previously found not to affect growth on cholesterol or β-sitosterol. To further elucidate their *in vivo* physiological role, we have cloned the *ltp3* and *ltp4* gene homologs of *Rhodococcus rhodochrous* DSM43269, and performed gene deletions in derived mutant strain RG32 which selectively degrades sterol side chains. Deletion of either *ltp3* or *ltp4* resulted in complete abolishment of degradation of the C24-branched side chain of β-sitosterol and campesterol, but not of the side chain of cholesterol, indicating a role of *ltp3* and *ltp4* in the removal of the C24-moiety. Bioinformatic analysis of $Ltp3_{DSM43269}$ and $Ltp4_{DSM43269}$ showed that the catalytic residues characteristic for thiolases are not conserved in their amino acid sequences. Furthermore, removal of the C24-branched side chain carbons of β-sitosterol was previously shown to proceed via aldolytic cleavage rather than by β-oxidation. Our results therefore suggest that *ltp3* and *ltp4* most likely encode aldol lyases, rather than thiolases. This is the first report on the identification and characterization of enzymes (Ltp3 and Ltp4) with specific and essential roles in carbon-carbon cleavage of C24-branched chain sterols in *Rhodococcus* strains, most likely acting as aldol lyases. The results provide a clear contribution to our understanding of sterol degradation in *Actinobacteria*, including pathogenic strains such as *Rhodococcus equi* and *Mycobacterium tuberculosis*.

INTRODUCTION

Actinobacteria, including those belonging to the genera of *Rhodococcus*, *Nocardia* and *Mycobacterium*, are renowned for their ability to metabolize a wide array of organic compounds, including steroids and sterols (Arima *et al*., 1969; Martin 1977; van der Geize and Dijkhuizen, 2004). Two oxidative pathways are involved in microbial sterol degradation: steroid ring oxidation and sterol side chain degradation (Dodson and Muir, 1958a, 1958b; Sih *et al*., 1968a, 1968b). The mechanism of microbial sterol side chain degradation has been elucidated at the biochemical level by studying cell-free systems of *Mycobacterium* and *Nocardia* strains, as well as by the identification of the intermediates formed (Sih *et al*., 1967, 1968a, 1968b; Martin and Wagner, 1976; Fujimoto *et al*., 1982a, 1982b). In addition, a cholesterol catabolic gene cluster of *R. jostii* RHA1 is comprised of a subset of genes, i.e. *ro04678-ro04695*, with a suggested role in sterol side chain degradation (van der Geize *et al*., 2007). Still, extremely limited information is available on the genes and enzymes involved in microbial sterol side chain degradation.

Sterol side chain degradation is initiated by C26-hydroxylation followed by complete oxidation to yield a sterol C26 carboxylic acid (Zaretskaya *et al*., 1968) (Fig. 1, compounds II and VII for cholesterol and C24-branched sterols, respectively). Recent studies showed that rhodococcal *cyp125* encodes a cytochrome P450 monooxygenase essential for formation of this sterol C26 carboxylic acid intermediate (Rosłoniec *et al*., 2009). Biochemical studies on the CYP125 homolog of *Mycobacterium tuberculosis* H37Rv revealed that $CYP125_{H37Rv}$ catalyzes hydroxylation of C26 or C27 (Capyk *et al*., 2009; McLean *et al*., 2009; Ouellet *et al*., 2010). Interestingly, *cyp125* of *M. tuberculosis* H37Rv is located adjacent to *fadA5* (*Rv3546*), which was recently characterized to encode a thiolase required for the formation of 4-androstene-3,17-dione (AD) and ADD by side chain degradation of cholesterol (Nesbitt *et al*., 2009). Upon formation of the acid intermediate, the sterol side chain is degraded via a mechanism similar to β-oxidation of fatty acids (Sih *et al*., 1967, 1968a, 1968b). This process is initiated by CoA activation of the carboxylic acid and mediated by a sterol-CoA ligase (Ambrus *et al.*, 1995). Recently, we have identified FadD19 of *Rhodococcus rhodochrous* DSM43269 as a sterol-CoA ligase with a specific role in the degradation of C24-branched chain sterols, e.g. β-sitosterol and campesterol (Fig. 1B) (Wilbrink *et al*., submitted [Chapter 4]). Following the formation of the CoA-activated intermediate, the side chain of cholesterol is shortened via C24-carboxylic acid (Fig. 1, compound III) and C22-carboxylic acid (Fig. 1, compound IV) intermediates, with concomitant release of propionyl-CoA and acetyl-CoA, respectively. Eventually, an aldolytic cleavage reaction is suggested to be responsible for formation of 17-keto-steroids, such as 1,4-androstadiene-3,17-dione (ADD, Fig. 1A, compound V) (Sih *et al*., 1967, 1968a, 1968b). Removal of the ethyl and methyl branches of β-sitosterol or campesterol, respectively, is initiated by carboxylation of the C28 carbon, followed by a hydration reaction, and ultimately release of propionyl-CoA via an aldolytic cleavage of the C24-C25 bond (Fig. 1B, Chen, 1985; Fujimoto *et al*., 1982b).

Aim of our work is to decipher the physiological roles of genes and enzymes involved in microbial sterol side chain degradation. For this purpose, *R. rhodochrous* DSM43269 strain

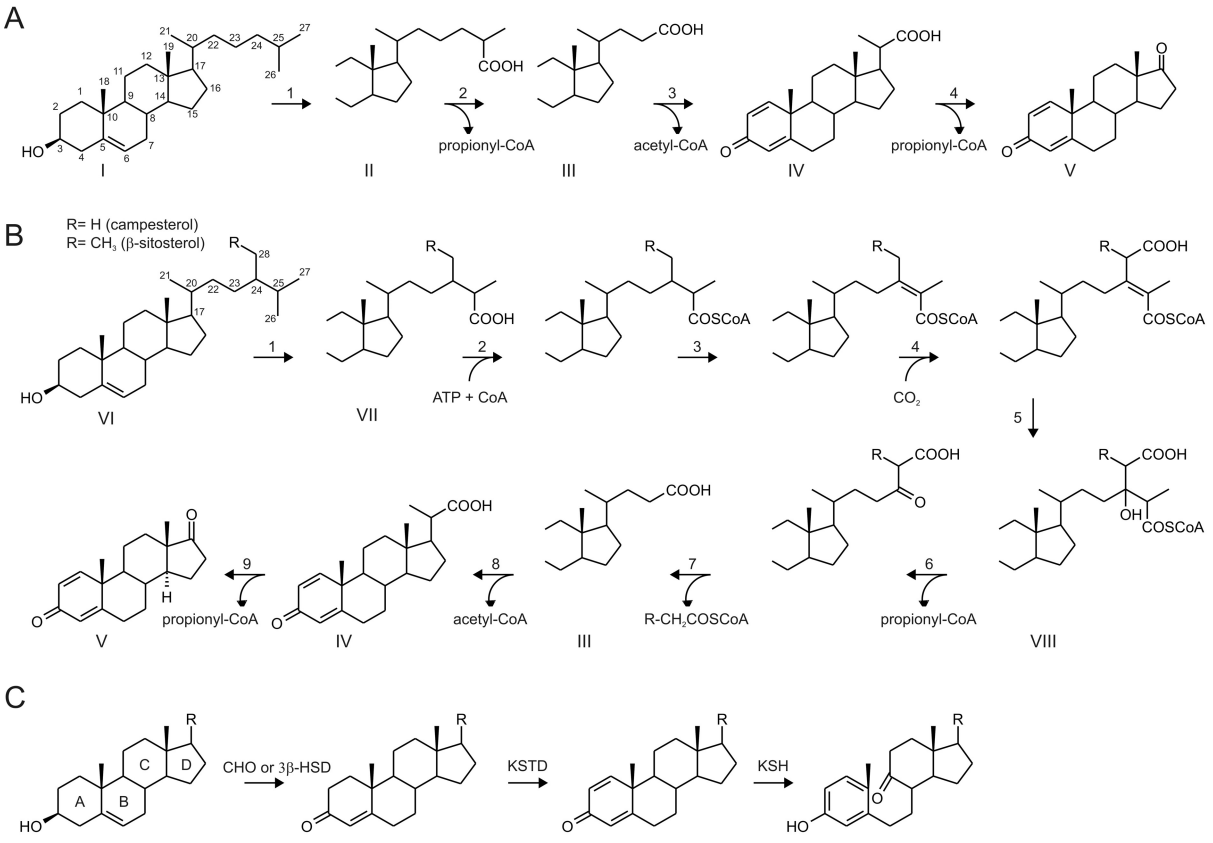


Fig. 1. Scheme showing the side chain degradation of (A) cholesterol and (B) the C24-branched sterols β-sitosterol and campesterol. Initial reactions involved in opening of the steroid nucleus (C).

RG32 has been constructed, lacking all 3-ketosteroid 9α-hydroxylase (KSH) activity (Fig. 1C) (Petrusma *et al.*, in preparation). Mutant strain RG32, completely blocked in steroid ring degradation, is capable of selective sterol side chain degradation, and accumulates ADD (Fig. 1A, compound V) and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC, Fig. 1A, compound IV) from various sterols (Rosłoniec *et al*., 2009). The strain RG32 phenotype thus is useful as a tool to verify the role of candidate genes in sterol side chain degradation. Here, we report on a mutational analysis of *ltp3* and *ltp4* in *R. rhodochrous* strain RG32 and propose physiological roles of these genes in sterol side chain degradation. *ltp3* (*ro04683*) and *ltp4* (*ro04684*) are part of the cholesterol catabolic gene cluster in RHA1 and were previously shown to be up-regulated during growth of *R. jostii* strain RHA1 on cholesterol (van der Geize *et al*., 2007). We also have investigated the role of the $fadA5_{DSM43269}$ homolog in cholesterol side chain degradation in *R. rhodochrous* by gene inactivation in strain RG32. The results provide a clear contribution to our understanding of sterol degradation in *Actinobacteria*, including pathogenic strains such as *Rhodococcus equi* and *M. tuberculosis*.

RESULTS

Cloning of *ltp3* and *ltp4* orthologs from *R. rhodochrous* DSM43269

R. rhodochrous RG32, devoid of all 3-ketosteroid 9α-hydroxylase (KSH) activity and thus blocked in steroid ring attack, is capable of selective sterol side chain degradation (Wilbrink

et al., submitted [Chapter 4]). To assess the roles of *ltp3* and *ltp4* in sterol side chain degradation we identified and cloned their orthologs in *R. rhodochrous* DSM43269 by screening a genomic library with degenerate primers based on highly conserved amino acid sequences present in actinobacterial Ltp3 homologs. An individual clone was identified carrying full-length $ltp3_{DSM43269}$ on a 5.7 kb DNA insert in addition to several other orthologs of RHA1 genes (Fig. 2). The genetic organization of these gene orthologs in DSM43269 was identical to that in strain RHA1 (Fig. 2). The proteins shared 74-92% amino acid sequence identities with their RHA1 orthologous counterparts.

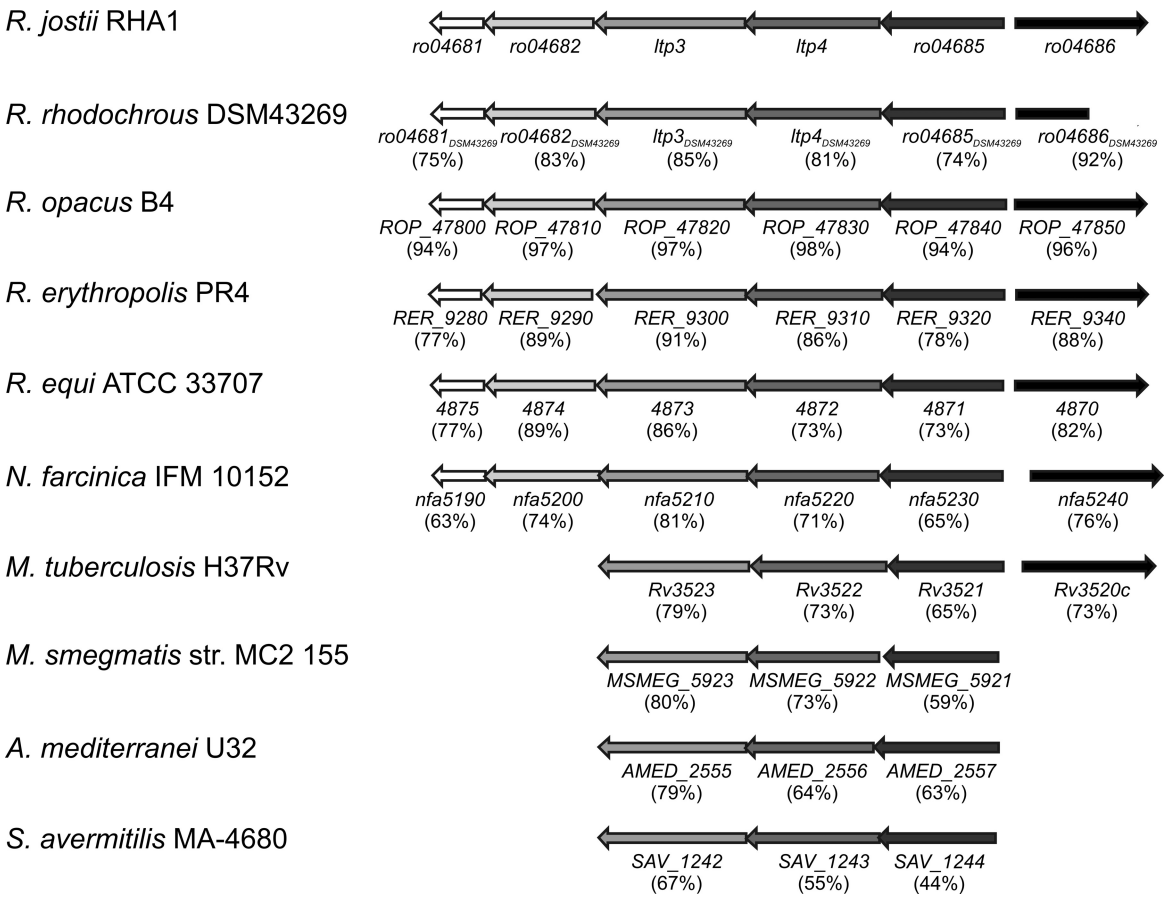


Fig. 2. Genetic organization of the *ro04681-ro04686* genes in *R. jostii* RHA1 and their homologs in other *Actinobacteria*. The respective amino acid sequence identities of the homologs to the RHA1 proteins are shown below the gene names.

Bioinformatic analysis of *R. rhodochrous* DSM43269 Ltp3 and Ltp4

Bioinformatic analysis revealed relatively low amino acid sequence identity between Ltp3 and Ltp4 (23%). Ltp3 and Ltp4 are conserved in many other sterol degrading *Actinobacteria* having 55-86% protein sequence identity (Fig. 2). In these *Actinobacteria*, *ltp3* and *ltp4* are organized in an apparent operon, consisting of at least three and up to five genes. Sequence comparison of $Ltp3_{DSM43269}$ and $Ltp4_{DSM43269}$ with characterized proteins from the Swiss-Prot database further showed that they both show highest similarity with eukaryotic sterol carrier protein x (SCPx) (26 and 24% identity, respectively). The N-terminal domain of SCPx

exhibits 3-ketoacyl-CoA thiolase activity specific for branched chain acyl-CoA substrates (Seedorf *et al*., 1994; Wanders *et al*., 1997). Sequence alignment of $Ltp3_{DSM43269}$ and $Ltp4_{DSM43269}$ with other well-studied thiolases revealed that the catalytic residues Cys, His and Cys (Cys_{125}, His_{375} and Cys_{403} in yeast thiolase), which are highly conserved in catabolic thiolases (Haapalainen *et al*., 2006), are absent in $Ltp3_{DSM43269}$ and $Ltp4_{DSM43269}$ (Fig. 3). The $Ltp3_{DSM43269}$ sequence furthermore did not contain any of the thiolase signature sequence motifs (PF00108 and PF02803 (Finn *et al*., 2008)) which include the catalytic residues, while in $Ltp4_{DSM43269}$ only the latter was semi-conserved. Strikingly, in a phylogenetic analysis of annotated thiolase proteins of strain RHA1 and characterized thiolases (Peretó *et al*., 2005), $Ltp3_{DSM43269}$ and $Ltp4_{DSM43269}$ cluster separately from most of the previously characterized thiolase enzymes from various sources (Fig. 4). These data strongly suggest that *ltp3* and *ltp4* do not encode thiolase activity, but another enzymatic function.

```
SPOT          116 tpfvalnrqcssgl  368 ggaialghplgctg  401 smcig
RPOT          124 vplsavnrqcssgl  380 ggaialghplgctg  416 smcig
thiolase I    121 vpirtvnrqcssgl  378 ggaiaighplgatg  415 smcig
yusK           81 vpaitvnrycssgl  340 ggaialghplgctg  375 tmcig
FadA           82 vpavtvnrlcgssm  336 ggaialghplgcsg  371 tmcig
HMOT           83 tpaltinrlcgsgf  345 ggaialghplggsg  380 sacig
thiolase II    90 victtinkvcaagm  354 ggavslghplgcsg  389 sicng
ZAT            80 atawgmnqlcgsgl  340 ggaiaighpigasg  375 tlcig
XCAT           83 ipawscqmicgsgl  346 ggavalghplgmsg  381 alcig
SCPx           85 ipiinvnnncatgs  348 ggliskghplgatg  389 qhnlg
FadA5          80 vgattidaqcgsaq  336 ggaialghpvgstg  371 smcag
Ltp3           78 kpmlrvhtagsvgg  329 ggvlc-snpigasg  369 ghayg
Ltp4           74 vppineshvemdaa  300 ggplt-gnpmfsgg  333 ahats
PF00108        79 vpaltvnkvcgsgl
Pf02803                          73 ggaialghplgasg  108 alcig
                        *                 *                *
```

Fig. 3 Alignment of Ltp3, Ltp4 and FadA5 of *R. rhodochrous* DSM43269 with characterized thiolases: SCPx, FadA, YusK, RPOT, thiolase I, thiolase II, XCAT, HMOT, and SPOT (*Saccharomyces cerevisiae* peroxisomal 3-oxoacyl-CoA thiolase, CAA37893) and ZAT (*Zoogloea ramigera* acetoacetyl-CoA transferase, AAA27706), as used by Peretó *et al*. (2005) and the N and C-terminal thiolase motifs PF00108 and PF020803, respectively (Finn *et al*., 2008). The highly conserved and catalytically important residues Cys, His and Cys (marked by an asterisk *) are conserved in $FadA5_{DSM43269}$ but are not present in $Ltp3_{DSM43269}$ nor $Ltp4_{DSM43269}$.

Both *ltp3* and *ltp4* of *R. rhodochrous* are required for β-sitosterol side chain degradation

Unmarked in-frame single gene deletion mutants of $ltp3_{DSM43269}$ and $ltp4_{DSM43269}$ were constructed in strain RG32, yielding mutant strains RG32Δ*ltp3* and RG32Δ*ltp4*. Conceivably, these *ltp3* and *ltp4* single gene deletion mutants cross-complement each others. Therefore, a double unmarked gene deletion mutant was also constructed, designated RG32Δ*ltp3* Δ*ltp4*. PCR analysis and nucleotide sequencing of these mutants confirmed the correct introduction of the gene deletions. In case of single mutants, the correct presence of the adjacent *ltp3* or *ltp4* gene was also confirmed by PCR analysis.

The mutants RG32Δ*ltp3* Δ*ltp4*, RG32Δ*ltp3* and RG32Δ*ltp4* were tested in whole cell biotransformation assays for their ability to convert the side chain of cholesterol and β-sitosterol. None of the mutants was affected in cholesterol side chain degradation, producing ADD and 1,4-BNC at similar levels as observed with strain RG32 (Table 1). However, both *ltp3* and *ltp4* were essential for the degradation of the β-sitosterol side chain, since none of the three mutants produced any 1,4-BNC and ADD (Table 1). Interestingly, all three RG32 mutants accumulated small amounts of an intermediate that was not observed in parent strain RG32. Mass spectrometry analysis of the intermediate revealed a base peak of m/z = 122, typical for a 3-keto-1,4-diene steroid ring structure (Brooks, 1979) and a molecular ion of m/z = 410. These values are in agreement with 1,4-sitostadiene-3-one (M_w = 410), which is the predicted product of A-ring oxidation of β-sitosterol through the combined action of cholesterol oxidase or 3β-hydroxysteroid dehydrogenase, and 3-ketosteroid Δ1-dehydrogenase activities (Fig. 1C) present in *Rhodococcus* strains (van der Geize *et al*., 2007; Knol *et al*., 2008; Rosłoniec *et al*., 2009). The accumulation of 1,4-sitostadiene-3-one further supported our conclusion that deletion of $ltp3_{DSM43269}$ and $ltp4_{DSM43269}$ in strain RG32 resulted in impaired β-sitosterol side chain degradation. Interestingly, the conversion of campesterol (Fig. 1) was also blocked in all three mutants (Table 1), suggesting that $ltp3_{DSM43269}$ and $ltp4_{DSM43269}$ are essential for C24-branched-chain sterol side chain degradation specifically.

In view of the organization of *ltp3* and *ltp4* in an apparent operon it was important to exclude the possibility that the phenotypes of strain RG32Δ*ltp3* and strain RG32Δ*ltp4* were caused by polar effects. Indeed, both RG32Δ*ltp3* and RG32Δ*ltp4* could be successfully complemented when the deleted gene was provided *in trans*, as shown by their regained ability to form ADD and 1,4-BNC from β-sitosterol in whole-cell biotransformations (Table 1). This conclusively demonstrated that the observed phenotypes were caused by deletion of the genes itself, rather than by polar effects.

Table 1. Whole-cell biotransformations of sterols by strain RG32 and mutants thereof. Conversion rates are depicted as the molar % of sterol conversion into ADD and 1,4-BNC, as measured by HPLC after 3 days of incubation. ND, not determined; -, below detection limit.

Strain	β-sitosterol		campesterol		cholesterol	
	ADD (%)	1,4-BNC (%)	ADD (%)	1,4-BNC (%)	ADD (%)	1,4-BNC (%)
RG32	7 ± (2)	67 ± (7)	1 ± (0)	48 ± (10)	3 (±1)	73 (±12)
RG32Δ*ltp3* Δ*ltp4*	-	-	-	-	5 (2)	71 (±8)
RG32Δ*ltp3*	-	-	-	-	4 (±1)	70 (±3)
RG32Δ*ltp4*	-	-	-	-	3 (±1)	71 (±2)
RG32Δ*ltp3* + pCOMP*ltp3*	4 ± (1)	22 ± (3)	ND	ND	ND	ND
RG32Δ*ltp4* + pCOMP*ltp4*	7 ± (1)	42 ± (5)	ND	ND	ND	ND
RG32Δ*ltp3* Δ*ltp4* Ω*fadA5*	-	-	ND	ND	3 (±2)	72 (±20)
RG32Ω*fadA5*	7 ± (1)	70 ± (8)	ND	ND	4 (±2)	74 (±9)

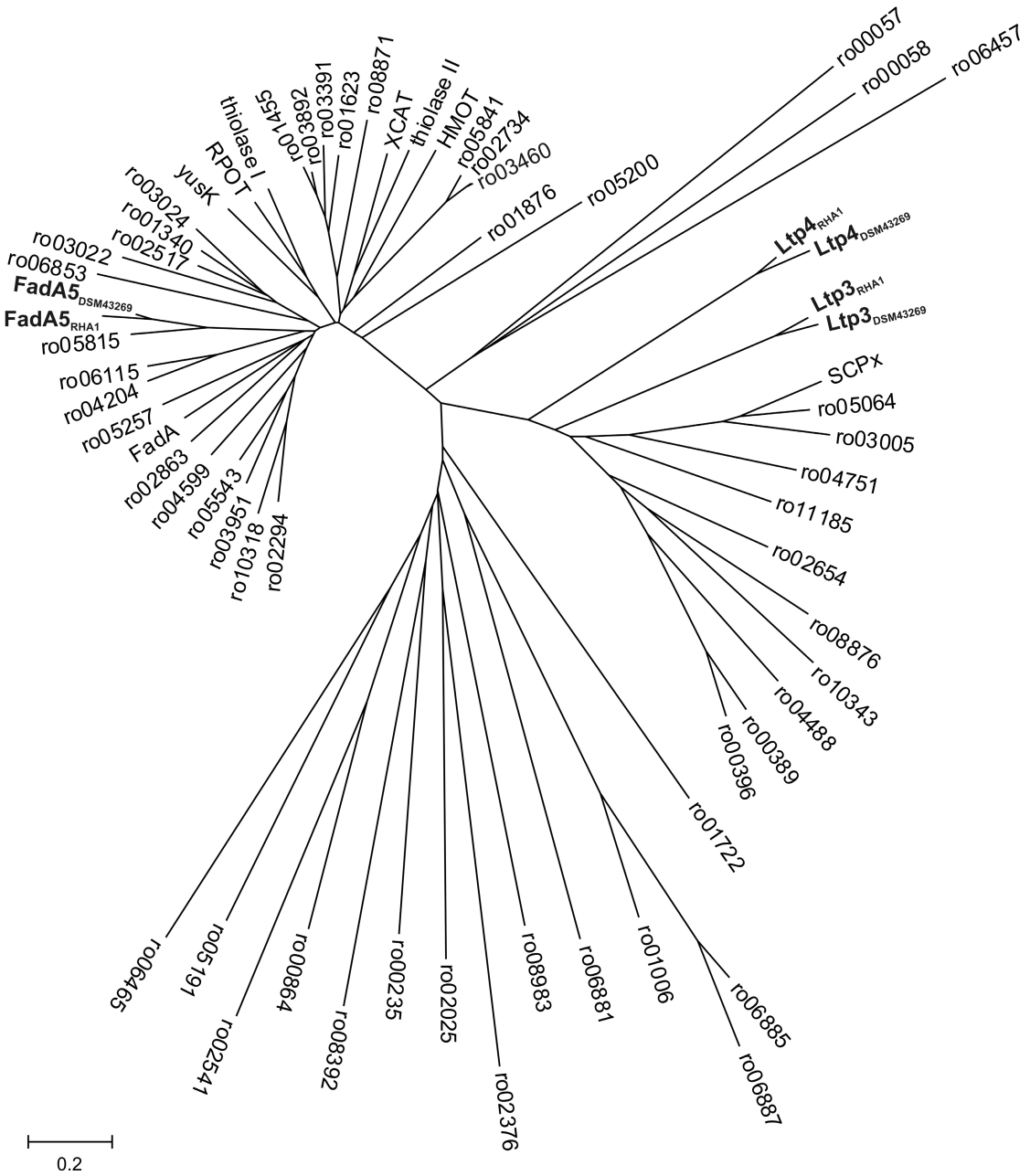


Fig. 4. Phylogenetic tree comprised of full length protein sequences annotated as thiolases (or thioesterases) in the *Rhodococcus jostii* RHA1 genome (www.rhodococcus.ca) and the characterized thiolases FadA, HMOT, RPOT, XCAT, thiolase I, thiolase II, YusK and SCPx (Peretó *et al*., 2005). Ltp3, Ltp4 and FadA5 of strains RHA1 and DSM43269 are shown in bold.

FadA5$_{DSM43269}$ is not essential for cholesterol degradation in *R. rhodochrous* RG32

FadA5$_{H37Rv}$ was previously identified as a 3-ketoacyl-CoA thiolase essential for side chain degradation of cholesterol in *M. tuberculosis* (Nesbitt *et al*., 2009). Bioinformatic analysis of a previously cloned genomic fragment of *R. rhodochrous* DSM43269 carrying *cyp125* (Rosłoniec *et al*., 2009) identified a FadA5$_{DSM43269}$ ortholog (accession number FJ824698) displaying 74% amino acid sequence identity with FadA5$_{H37Rv}$. To substantiate a role of *fadA5$_{DSM43269}$* in sterol side chain degradation in *R. rhodochrous* DSM43269, this gene was inactivated in strains RG32 and RG32Δ*ltp3* Δ*ltp4* and sterol degradation ability was assessed.

In both *fadA5*$_{DSM43269}$ inactivated mutants, however, cholesterol side chain degradation was still unimpaired (Table 1). Moreover, β-sitosterol side chain degradation was not impaired in RG32Ω*fadA5*$_{DSM43269}$ (Table 1). These results indicate that, contrary to *M. tuberculosis*, *fadA5* in *R. rhodochrous* DSM43269 is not essential for side chain degradation of cholesterol or β-sitosterol. Further bioinformatic analysis of the *R. jostii* RHA1 genome revealed that *ro05815* encodes a protein with high amino acid similarity with FadA5$_{RHA1}$ and FadA5$_{H37Rv}$ (67% and 64 %, respectively). Interestingly, reciprocal best hits of Ro05815 could be identified in all available *Rhodococcus* genomes with high protein sequence similarities (82-92%), but not in *M. tuberculosis* H37Rv. Analogously, *R. rhodochrous* DSM43269, whose genome sequence is not known, may also contain such gene homolog, thus explaining the lack of phenotype in the *fadA5*$_{DSM43269}$ inactivated mutants. Alternatively, the FadA5 homologs of H37Rv and DSM43269 may have different enzymatic activities.

DISCUSSION

The current study reports on the molecular characterization of *ltp3* and *ltp4* in *R. rhodochrous* DSM43269. Ltp3$_{DSM43269}$ and Ltp4$_{DSM43269}$ are found conserved in several other *Actinobacteria*, displaying high amino acid sequence similarity with their counterparts in *R. jostii* strain RHA1 (85 and 81%, respectively). Previous work on strain RHA1 suggested that *ltp3*$_{RHA1}$ and *ltp4*$_{RHA1}$ may have a role in side chain degradation of cholesterol. The *ltp3*$_{RHA1}$ and *ltp4*$_{RHA1}$ are located within the cholesterol catabolic gene cluster (van der Geize *et al.*, 2007), in-between *cyp125*$_{RHA1}$, encoding a steroid C26-hydroxylase (Fig. 1A, step 1; Rosłoniec *et al*., 2009), and *fadD19*$_{RHA1}$, encoding a steroid CoA ligase (Fig. 1A, step 2; Wilbrink *et al*., submitted, [Chapter 4]). Also, transcriptional analysis showed that expression of *ltp3*$_{RHA1}$ was up-regulated during growth of strain RHA1 on cholesterol (van der Geize *et al*., 2007). A mutant strain RG32 of *R. rhodochrous* DSM43269, devoid of all KSH activity and capable of selective sterol side chain degradation, was previously constructed and shown to be a suitable tool to establish the role of candidate genes in sterol side chain degradation (Petrusma *et al*., in preparation; Wilbrink *et al*., submitted, [Chapter 4]). Surprisingly, deletion of either *ltp3*$_{DSM43269}$ or *ltp4*$_{DSM43269}$ in strain RG32 did not affect cholesterol side chain degradation. Consistent with this finding we previously found that deletion of a suite of genes that included *ltp3* and *ltp4* in strain RHA1 did not abolish growth on cholesterol (Wilbrink *et al*., submitted). Interestingly, side chain degradation of β-sitosterol and campesterol was found to be complete blocked in RG32Δ*ltp3* and RG32Δ*ltp4* (Table 1). β-Sitosterol and campesterol are phytosterols with C24β-ethyl and C24α-methyl branched side chains, respectively, that distinguish them from cholesterol (Fig. 1). Thus, *ltp3* and *ltp4* appear to be essential for the degradation of C24-branched sterol side chains specifically.

Amino acid sequence comparison of Ltp3$_{DSM43269}$ and Ltp4$_{DSM43269}$ with characterized thiolases further revealed that none of the typical catalytic residues or Pfam sequence motifs for thiolases were conserved (Fig. 3), suggesting that *ltp3* and *ltp4* do not encode typical thiolase activities, but another enzymatic function. Fujimoto *et al*. (1982a) indeed

suggested that compound VIII (Fig. 1B) is cleaved by an aldol lyase similarly to the cleavage of other tertiary β-hydroxy CoA esters, such as (β-methyl-) malyl-CoA (Hacking and Quayle, 1974) and hydroxymethylglutaryl-CoA (Stegink and Coon, 1968). Hence, we hypothesize that *ltp3* and *ltp4* do not encode thiolase enzymes but rather encode aldolytic enzymes involved in the cleavage of the C-C bond at position 24-25 of C24-branched chain sterols (Fig. 1A, step 6). This activity would also explain why Ltp3 and Ltp4 do not appear to have a role in the degradation of the side chain of cholesterol, since such aldol lyase activity is not needed for the cleavage of the C24-C25 bond of cholesterol during side chain degradation (Fig. 1A). Our hypothesis is further supported by the observation that Ltp3 and Ltp4 proteins phylogenetically cluster with SCPx separate from previously characterized thiolases (Fig. 4). SCPx is a eukaryotic enzyme with branched chain 3-ketoacyl-CoA thiolase activity (Seedorf *et al*., 1994; Wanders *et al*., 1997). Interestingly, SCPx is involved in steroid metabolism performing the last step of β-oxidation of 3α,7α,12α-trihydroxy-5β-cholest-24-ene-26-oyl-CoA and 3α,7α-dihydroxy-5β-cholest-24-ene-26-oyl-CoA thereby forming their 24-oxo products in the process of cholic acid formation from cholesterol (Antonenkov *et al*., 1997; Bun-ya *et al*., 1998; Takeuchi *et al*., 2004).

It currently remains unclear from our work whether Ltp3 and Ltp4 form homo- or heteromeric complexes. Active malyl-CoA lyase enzyme of *Methylobacterium extorquens* (formerly *Pseudomonas* AM1) was shown to consist of a homohexameric protein of 190 kDa (Hacking and Quayle, 1974). Additional biochemical studies are required to definitively establish the exact functions of Ltp3 and Ltp4 in sterol side chain degradation. However, such studies are hampered by the lack of availability of the required sterol substrates (Fig. 1B, compound VIII) for the reaction(s) expected to be catalyzed by these enzymes (Fig. 1, step 6).

In our search for genes essential for cholesterol side chain degradation by β-oxidation we decided to target inactivation of $fadA5_{DSM43269}$, since FadA5 of *M. tuberculosis* H37Rv was characterized as a 3-ketoacyl-coA thiolase essential for cholesterol side chain degradation (Nesbitt *et al*., 2009). In contrast to *M. tuberculosis* H37Rv, however, cholesterol degradation by both RG32 and RG32Δ*ltp3* Δ*ltp4* with inactivated $fadA5_{DSM43269}$ was unimpaired. Conceivably, the genome of *R. rhodochrous* encodes one or more other thiolases that provides enzymatic compensation for FadA5 activity. This hypothesis is supported by the observation that the genome of the closely related *R. jostii* RHA1 contains a close homolog of *fadA5*, i.e. *ro05815* (Fig. 4), which is conserved in all available *Rhodococcus* genome sequences, but has no counterpart in *M. tuberculosis* H37Rv. Analogously, the genome of *R. rhodochrous* DSM43269 may encode the same gene homolog. However, since the complete genome sequence of *R. rhodochrous* DSM43269 is not known and the exact function of *ro05815* has not been studied, the possibility that the FadA5 enzymes from *R. rhodochrous* DSM43269 and *M. tuberculosis* H37Rv catalyze different reactions, and are involved in different physiological processes, cannot be ruled out.

This is the first report on the identification and characterization of enzymes (Ltp3 and Ltp4) with specific and essential roles in carbon-carbon cleavage of branched chain sterols in *Rhodococcus* strains, most likely acting as aldol lyases. The results provide a clear contribution to our understanding of sterol degradation in *Actinobacteria*, including pathogenic strains such as *R. equi* and *M. tuberculosis*.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

Rhodococcus rhodochrous DSM43269 (and mutants derived) were grown in Luria-Bertani (LB, Sigma, Zwijndrecht, the Netherlands) at 30°C and liquid cultures were shaken at 220 rpm. *Escherichia coli* strains were grown in LB medium at 37°C and shaking at 220 rpm for liquid cultures. For growth on solid medium, 1.5% (w/v) Difco™ Agar (Becton, LePont de Claix, France) was added. *E. coli* DH5α and pBlueScript(II)KS were used for gene cloning. *E. coli* S17-1 was used for conjugation of pK18*mobsac*B derivatives to *Rhodococcus*. All conjugational steps were plated on LB-agar at 30°C, sucrose sensitivity of *Rhodococcus* was tested on LB-agar supplemented with 10% (w/v) sucrose. Kanamycin was used at a concentration of 25 µg/ml. Cholesterol (>95%, ash-free) and campesterol (65%) were obtained from Sigma (Schnelldorf, Germany). β-Sitosterol was purchased from Acros (Geel, Belgium) and was a mixture of β-sitosterol (75-80%), β-sitostanol (10-14%) and campesterol (6-9%).

General cloning techniques

Procedures for the manipulation and analysis of DNA were performed essentially as described by Sambrook and co-workers (1989). DNA modifying enzymes (restriction enzymes, T4 DNA ligase and DNA polymerases) were purchased from Roche (Mannheim, Germany), New England Biolabs (Beverly, Mass. USA) or Fermentas (St. Leon-Rot, Germany) and were used according to the manufacturer's protocol.

All PCR products were obtained using *Pwo* polymerase (Roche) or *Pfu* polymerase (Fermentas) under standard conditions: denaturation at 94°C for 2 min followed by 30 cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 2 min. DNA fragments were purified from agarose gel using Sigma Genelute™ gel extraction kit according to the manufacturers protocol.

Screening of a R. rhodochrous DSM43269 genomic library for ltp3 ortholog

A genomic library of *R. rhodochrous* DSM43269 in pRESQ (Petrusma *et al*., 2009) was screened for the presence of a *ltp3* orthologs using gene specific degenerate primers based on conserved amino acids of Ltp3 (supplemental Table S1). The genomic library of *R. rhodochrous* DSM43269 was transferred to *E. coli* DH5α by transformation and plated on LB with 25 µg ml^{-1} kanamycin. All colonies on a plate were resuspended in liquid LB medium, the total plasmid pool was isolated, and checked with PCR using the degenerate primers for

the presence of the target gene. The procedure of transformation, plating dilutions and PCR was repeated until a single clone was identified, named pRESQ4683, containing *ltp3*$_{DSM43269}$.

Construction of targeted gene disruptions and unmarked gene deletions in R. rhodochrous RG32

Unmarked gene deletion mutants of *R. rhodochrous* RG32 were obtained as described previously (van der Geize *et al*., 2001). Plasmid constructs to perform targeted gene deletions of *ltp3*$_{DSM43269}$ and *ltp4*$_{DSM43269}$ were constructed by PCR amplification of up- and downstream regions of the target genes using the primers listed in supplemental Table S1. The obtained amplicons were separately subcloned into *Eco*RV digested pBlueScript(II)KS. The amplicons were then joined by ligation upon *Bgl*II/*Hind*III (Δ*ltp3*) or *Hind*III/*Nde*I (Δ*ltp4* and Δ*ltp3* Δ*ltp4*) restrictions. The deletion cassettes were then cloned into pK18*mobsac*B using the *Xba*I and *Hind*III restriction sites.

Gene deletions were confirmed by PCR with specific primers (supplemental Table S1) using chromosomal DNA isolated from the mutants.

Complementation of mutants RG32Δltp3 and RG32Δltp4

Chromosomal DNA of *R. rhodochrous* DSM43269 was used to amplify *ltp3* and *ltp4* using Pwo DNA polymerase (Roche) using the primers listed in supplemental Table S1. The obtained PCR products of *ltp3*$_{DSM43269}$ (1.2 kb) and *ltp4*$_{DSM43269}$ (1.1 kb) were digested with *Acc*65I and cloned behind the *aphII* promoter region of *Eco*RV/*Acc*65I digested pBs-Pkan (van der Geize *et al*., 2008a), yielding pBs-Pkan*ltp3* and pBs-Pkan*ltp4*, respectively. The Pkan-*ltp3* and Pkan-*ltp4* cassettes were then cloned into *Spe*I/*Acc*65I digested *Rhodococcus*-*E. coli* shuttle vector pRESQ. The resulting constructs were named pCOMP*ltp3* and pCOMP*ltp4*, respectively. The plasmids were mobilized to the respective mutants by electro-transformation as described by Fernandes *et al*. (2001).

HPLC and GC-MS analysis of sterol bioconversions

For bioconversion experiments sterols were dissolved in acetone (25 mM) and added at a final concentration of 0.5 mM to mid-exponential phase grown *Rhodococcus* cultures which were incubated for another three days at 30°C with shaking (220 rpm). For HPLC analysis, 1 ml of cell culture was centrifuged (1 min., 16,000 g), the supernatant was diluted five fold with methanol (100%) and filtered, prior to injection. Steroids were separated on an Alltima C18 column (250 mm x 4.6 mm, 5 µm) at 35°C, using a mobile phase consisting of methanol:water:formic acid (80:19:1) at a flow rate of 1 ml/min and UV detection (254 nm). ADD and 1,4-BNC standards were used to quantify conversion rates as described previously (Wilbrink *et al*., submitted).

For GC-MS analysis, samples were prepared by adding 50 µl of H_2SO_4 (10%, v/v) to 0.5 ml of cell culture followed by extraction with two volumes of ethyl acetate and vigorous shaking. The organic fraction was collected and dried under a stream of nitrogen and derivatized using trimethylchlorosilane (TMCS) as described (Song *et al*., 2003).

Phylogenetic tree construction

The amino acid sequences of thiolases encoded by strain RHA1 were obtained from the RHA1 genome website (www.rhodococcus.ca) by database searches for genes annotated as thiolase or thioesterase, or with a Cluster of Orthologous Groups (COG) prediction as acetyl-CoA acetyltransferase. The obtained full-length protein sequences were subsequently aligned using ClustalW (Thompson *et al*., 1994). Phylogenetic and molecular evolutionary analyses were performed using MEGA version 4 (Tamura *et al*., 2007). The following characterized thiolases (Peretó *et al*., 2005) from various organisms were retrieved from GenBank; *E. coli* FadA (AAC76848), *Homo sapiens* mitochondrial 3-oxoacyl thiolase, HMOT (BAA03800), *Rattus norvegicus* peroxisomal thiolase I, RPOT (BAA14106), *Xenopus laevis* cytosolic thiolase II, XCAT (AAD34967), *Arabidopsis thaliana* thiolase I (NP_171965) and thiolase II (NP_199583), *Bacillus subtilis* YusK (BG14023), and *H. sapiens* sterol carrier protein 2, SCPx (NP_002970).

DDBJ/EMBL/GENBANK DATABASE ACCESSION NUMBERS

The *R. rhodochrous* DSM43269 sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession number HQ184439.

ACKNOWLEDGEMENTS

This work was financially supported by the Integration of Biosynthesis and Organic Synthesis Program of Advanced Chemical Technologies for Sustainability. M. H. W. received a traineeship from the European Graduate College program of the Groningen Biomolecular Sciences and Biotechnology Institute.

REFERENCES

References are listed on pages 119-140.

CHAPTER 6

Summary and concluding remarks

Nederlandse samenvatting

References

SUMMARY AND CONCLUDING REMARKS

Sterols are abundantly present in Nature due to their omnipresence in eukaryotic cell membranes, e.g. in decaying plant material. Some bacteria, especially *Actinobacteria*, have evolved enzymatic pathways enabling them to completely degrade available sterols as carbon and energy sources to sustain growth (Chapter 1).
Bacterial degradation of sterols involves two oxidative pathways, i.e. the removal of the 17-alkyl side chain and degradation of the steroid nucleus. Selective sterol side chain degradation is of interest to the pharmaceutical industry as it results in the formation of steroid intermediates that can serve as building blocks for the synthesis of bioactive steroids used in medicine. Specific strains of *Mycobacterium*, *Nocardia* and *Rhodococcus* have been isolated that are able to selectively degrade sterol side chains because of (naturally or mutationally) impaired steroid ring degradation. The mechanism of microbial sterol side chain degradation is known, but detailed information of the genes and enzymes involved in this process is still lacking. This is largely due to the genetic inaccessibility of most of the strains studied and their employment of multiple genes encoding similar catabolic activities. The limited knowledge on bacterial sterol side chain degradation at the genetic and enzymatic levels hampers the rational construction and engineering of industrially important production strains with enhanced overall sterol transformation efficiency and specificity in formation of novel intermediate steroid structures (Chapter 1).

The research reported in this PhD thesis aimed to obtain detailed information on the genetics, biochemistry and physiology of bacterial sterol side chain degradation. The research focused on two genetically accessible nocardioform *Actinobacteria*, *Rhodococcus rhodochrous* strain DSM43269 and *Rhodococcus jostii* strain RHA1, both degrading sterol side chains. The complete genome sequence of *R. jostii* RHA1 (9.7 Mb) had become available in recent years enabling us to perform a transcriptomic analysis of steroid catabolism. In Chapter 2, the identification of a cholesterol catabolic gene cluster in *R. jostii* strain RHA1 by transcriptomic analysis is described. The cluster consists of 51 upregulated genes that occur in a 235 kb stretch of chromosomal DNA. Characterization of a subset of these genes, predicted to be involved in sterol side chain degradation, proved difficult in strain RHA1 due to its high steroid ring degradation activities. Attempts to completely inactivate steroid ring degradation activities in strain RHA1, aiming to construct a mutant strain selectively degrading sterol side chains only, failed. Therefore another strain was chosen as model organism to study sterol side chain degradation, i.e. *R. rhodochrous* DSM43269. Wild type strain *R. rhodochrous* DSM43269 (= IFO3338), a potent sterol and steroid degrader, had previously been shown to be capable of selective sterol side chain degradation when ketosteroid hydroxylase (KSH) activity, an enzyme essential for steroid ring opening, was chemically inhibited by iron chelators, resulting in accumulation of 1,4-androstadiene-3,17-dione (ADD) and 3-oxo-23,24-bisnor-1,4-choladiene-22-oic acid (1,4-BNC) from sterol substrates (Arima *et al*., 1978). Five *kshA* genes, encoding KSH activities,

subsequently were identified in strain DSM43269 (Petrusma *et al*, in preparation) and consecutively deleted, resulting in mutant strain RG32. Indeed, strain RG32 transforms the side chains of various sterols via β-oxidation, yielding 1,4-BNC (50-80% mole/mole) and ADD (1-7% mole/mole) (Chapter 4; Fig. 1). The incomplete transformation of sterols into ADD by strain RG32 may be attributed to cytotoxicity of ADD, as has been reported for *Mycobacterium* strains (Lee and Liu, 1992, Smith *et al*., 1993). Other explanations are that product inhibition of the enzymes responsible for degradation of 1,4-BNC occurs, and/or that the genes encoding these activities are repressed by negative feedback regulation. The removal of the short side chain of 1,4-BNC, leading to the formation of ADD, is believed to occur by aldolytic cleavage in *Nocardia* sp. (Sih *et al*., 1968a), but the enzyme(s) involved are currently not known. Alternatively, the side chain of 1,4-BNC can be decarboxylated yielding progesterone, as was shown in *Mycobacterium aurum* (Prome *et al*., 1983), followed by reactions catalyzed by a steroid Baeyer-Villiger monooxygenase and an esterase to produce the 17-hydroxysteroid testosterone (Miyamoto *et al*., 1995).
The strain RG32 phenotype enabled us to study the role of genes in sterol side chain degradation specifically: inactivation of genes encoding an essential step in this process will result in loss of ADD (and 1,4-BNC) formation, depending on the enzymatic step that is inactivated (see Chapters 3, 4 and 5 of this thesis). However, when multiple genes code for the same enzymatic activity, a phenotype may only be observed in mutants carrying multiple gene deletions. The construction of such multiple gene deletions in strain RG32 is possible using the *sacB* counter selection system (van der Geize *et al*., 2001).
Further characterization of strain RG32 and its sterol side chain degradation mechanism and regulation may eventually lead to an engineered strain with enhanced overall transformation of sterols into ADD or 1,4-BNC, and possible applications of this strain in sustainable large-scale production of steroids from phytosterols.

This thesis describes a detailed characterization of four sterol side chain degradation genes in strain RG32. These four genes, i.e. *cyp125*, *fadD19*, *ltp3* and *ltp4*, are located proximal to each other within the cholesterol gene cluster at a locus homologous to the one predicted to encode sterol side chain degradation in strain RHA1. Our data revealed that *cyp125* encodes a cytochrome P450 sterol C26-hydroxylase enzyme (Chapter 3), that *fadD19* is a steroid-CoA ligase required for C24-branched chain sterol side chain degradation (Chapter 4), and that the *ltp3* and *ltp4* genes most likely encode aldol lyase enzymes essential for C24-branched chain sterol side chain degradation (Chapter 5). Interestingly, especially degradation of sterols with C24-branched side chains (i.e. β-sitosterol, campesterol) was blocked in RG32 mutants inactivated in these enzymes.

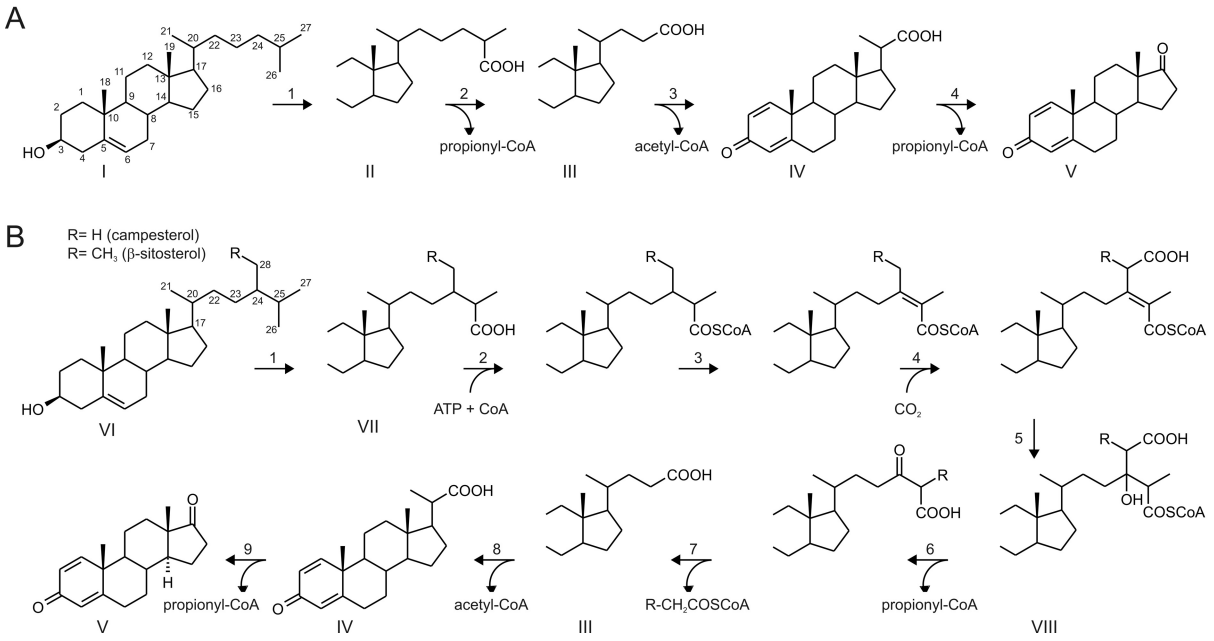


Fig. 1. Schematic overview of the side chain degradation pathways of (A) cholesterol and (B) the C24-branched chain sterols β-sitosterol and campesterol. CYP125 was characterized as a steroid 26-hydroxylase involved in step 1 (Fig. 1A and 1B). Furthermore, CoA ligase FadD19 is a steroid-CoA ligase, essential for C24-branched sterol side chain degradation and catalyzing step 2 (Fig. 1B). *In vitro*, FadD19 also catalyzes activation of 5-cholestene-26-oic acid-3β-ol (and metabolites thereof). Ltp3 and Ltp4 are both essential for degradation of C24-branched sterol side chains, most likely performing an aldol lyase reaction (step 6, Fig. 1B).

THE CHOLESTEROL CATABOLIC GENE CLUSTER OF *RHODOCOCCUS JOSTII* RHA1

Transcriptomic analysis of *R. jostii* RHA1 cells grown on cholesterol resulted in the identification of a cholesterol catabolic gene cluster encoded by *ro04482-ro04705* (Chapter 2). The gene cluster encodes proteins with high similarity to the characterized steroid degradation enzymes 3-ketosteroid Δ1-dehydrogenase (KSTD) and KSH, involved in the formation of 3-hydroxy-9,10-secoandrosta-1,3,5,(10)-triene-9,17-dione (3-HSA) during degradation of 4-androstene-3,17-dione (AD). Furthermore, seven *hsa* genes are present whose gene products were homologous to enzymes responsible for degradation of 3-HSA in related *Actinobacteria* and in other steroid degrading bacteria. Some of the encoded proteins were annotated as being involved in degradation of polyaromatic compounds (biphenyls), but biochemical studies of $HsaC_{H37Rv}$ and $HsaD_{H37Rv}$ have shown much higher specific activities (~ 40 fold) of these enzymes with steroids rather than with biphenyl metabolites (Chapter 2). The cholesterol catabolic gene cluster also encodes an operon (*mce4A-I*) responsible for cholesterol uptake, which is essential for growth on sterol substrates, but not steroids (van der Geize *et al*., 2007, [Chapter 2]; Mohn *et al*., 2008). Additionally, a suite of genes was identified comprising *ro04678-ro04695* (Fig. 2), which encodes putative β-oxidation enzymes suggested to be involved in 17-alkyl side chain degradation of sterols (Chapter 2). Since genetic information on sterol side chain degradation was virtually absent, these genes could not be compared with known sterol

side chain degradation genes and more detailed investigations thus were required to assess their exact role in sterol catabolism. The studies described in Chapters 3-5 unraveled the physiological functions of *cyp125* (= *ro04679*), *fadD19* (= *ro04689*), *ltp3* (= *ro04683*) and *ltp4* (= *ro04684*), respectively, all located within the sterol side chain degradation locus.

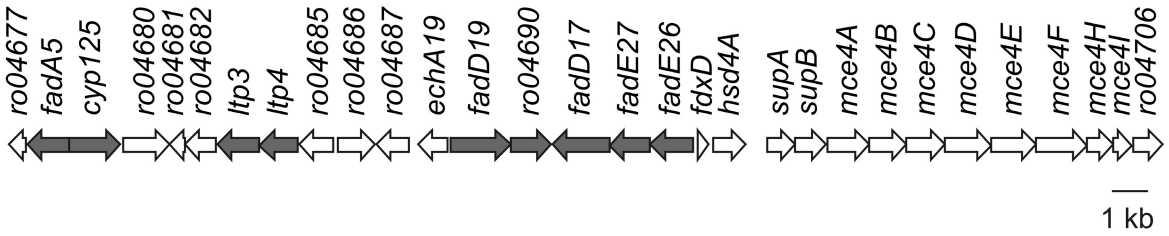


Fig. 2. Genetic organization of the sterol side chain degradation locus, *ro04677-hsd4A* and the downstream *mce4* operon, comprising *supA-ro04706*, encoding cholesterol transport in *R. jostii* strain RHA1 (McLeod *et al*., 2006; van der Geize *et al*., 2007; Mohn *et al*., 2008). Gene arrows indicated in grey were molecularly characterized in this thesis.

Interestingly, we further discovered that the cholesterol catabolic gene cluster was conserved in *Mycobacterium tuberculosis* H37Rv, a human pathogen and the causative agent of tuberculosis (TB). The essential roles of many of the genes encoded by this cluster in *M. tuberculosis* macrophage survival was described previously, but the metabolic role of these genes had remained unknown (Schnappinger *et al*., 2003; Rengarajan *et al*., 2005). Analogous to the role of the cholesterol catabolic genes in *Rhodococcus*, we predicted that the *M. tuberculosis* genes also function in cholesterol metabolism (van der Geize *et al*., 2007, [Chapter 2]). In subsequent studies, others have shown that *M. tuberculosis* is indeed able to metabolize cholesterol and that this feature is essential for its pathogenicity (Brzostek *et al*., 2007, 2009; Chang *et al*., 2009; Hu *et al*., 2009; Lack *et al*., 2009; Nesbitt *et al*., 2010; Pandey and Sassetti, 2008; Yam *et al*., 2009).

CYP125 IS A STEROID C26-HYDROXYLASE ESSENTIAL FOR INITIATION OF STEROL SIDE CHAIN DEGRADATION

The *cyp125* (= *ro04679*) gene of strain RHA1, encoding a cytochrome P450 monooxygenase, was highly upregulated during growth of this strain on cholesterol (Chapter 2). Deletion of *cyp125* in strain RHA1 resulted in impaired growth on 3-hydroxysterols, but not 3-keto derivatives (Rosłoniec *et al*., 2009, [Chapter 3]). Furthermore, heterologously expressed and purified $CYP125_{RHA1}$ was shown to bind sterols, however, no hydroxylase activity towards sterol substrates was detected using either chemically reduced enzyme or in the presence of various protein electron donors, such as KshB and the spinach ferredoxin and ferredoxin-reductase electron transport chain (Chapter 3). Possibly, these electron donor proteins are not compatible with $CYP125_{RHA1}$. The physiological redox proteins of $CYP125_{RHA1}$ thus remain to be identified, since their coding genes are not located proximal to *cyp125* in the strain RHA1 genome.

To verify our hypothesis that *cyp125* is involved in sterol side chain degradation we cloned the *cyp125* homolog of *R. rhodochrous* DSM43269 and inactivated the gene in derived strain RG32. Gene disruption mutant RG32Ω*cyp125* was shown to be completely blocked in side chain degradation of cholesterol and β-sitosterol (Chapter 3). Functional complementation experiments have been performed to ensure that the observed phenotype was solely due to inactivation of $cyp125_{DSM43269}$. Since mutant RG32Ω*cyp125* was able to transform the side chain of the C26-oic acid of cholesterol (i.e. 5-cholestene-26-oic acid-3β-ol), which is an expected side chain degradation metabolite (Fig. 1), it was concluded that CYP125 is involved in formation of the C26-oic acid of sterol, probably acting as a C26-hydroxylase (Chapter 3). Later biochemical studies on the CYP125 homolog of *M. tuberculosis* H37Rv have confirmed that CYP125 is capable of hydroxylation of sterol side chains at the C26 or C27 position (Capyk *et al*., 2009; McLean *et al*., 2009) and catalyzes the further oxidation to the C26/27-oic acid (Ouellet *et al*., 2010; Johnston *et al*., 2010). From these studies, it remained unclear whether CYP125 also catalyzes the subsequent oxidations to the C26-oic acid or that (an)other enzyme(s) catalyze these reactions.

Currently, the native redox partner proteins of rhodococcal CYP125 are not known. $CYP125_{H37RV}$ enzyme activity was observed when KshB (Rv3571) was used as an electron donor protein. Consistently, a *kshB* deletion mutant in *Rhodococcus erythropolis* SQ1 (i.e. strain RG4) was impaired in phytosterol degradation; KshB therefore was proposed to be part of the sterol C26-hydroxylase system, in addition to its role in KSH activity (van der Geize *et al*., 2002b). In future work, the various rhodococcal KshB variants available could be tested as reductase partner for rhodococcal CYP125. The side chain of the sterol C26-oic acid formed by CYP125 activity is further degraded by a mechanism similar to β-oxidation of fatty acids.

$FADD19_{DSM43269}$ IS A STEROID-COA LIGASE ESSENTIAL FOR DEGRADATION OF C24-BRANCHED STEROLS

Initiation of the β-oxidation cycle occurs by CoA-activation of the terminal carboxylic acid, catalyzed by CoA ligase enzymes. The sterol side chain degradation locus of strain RHA1 comprises two genes annotated as CoA ligases, *fadD17* and *fadD19,* that were both upregulated during growth on cholesterol (Chapter 2). To establish their exact role in sterol catabolism, their gene homologs were cloned from *R. rhodochrous* DSM43269 and inactivated in derived strain RG32. Interestingly, inactivation of these genes alone (i.e. deletion mutant RG32Δ*fadD19* and disruption mutant RG32Ω*fadD17*), or simultaneously (i.e. mutant RG32Δ*fadD19* Ω*fadD17*) in strain RG32 did not affect cholesterol side chain degradation. On the other hand, mutant strain RG32Δ*fadD19* was impaired in degradation of the C24-branched side chains of β-sitosterol and campesterol (Chapter 4). To confirm that $FadD19_{DSM43269}$ is indeed a steroid-CoA ligase the $fadD19_{DSM43269}$ gene was cloned and heterologously expressed in *E. coli*; the recombinant protein produced was shown to be active towards 5-cholestene-26-oic acid-3β-ol (and its 3-keto-4-ene derivative). Unfortunately, steroids possessing a C24-branched side chain with a C26-oic acid

functionality, the expected native substrates of FadD19$_{DSM43269}$, were not available and hence could not be tested. The mutant phenotype of FadD19, however, showed that it is essential for β-sitosterol and campesterol side chain degradation. From our results it remains elusive whether side chain activation during cholesterol catabolism is performed by multiple CoA ligases (including FadD19), or that a yet unidentified, highly specific enzyme catalyzes this reaction *in vivo*. Since FadD19$_{DSM43269}$ is able to activate the side chains of 5-cholestene-26-oic acid-3β-ol metabolites *in vitro*, and both *fadD19$_{RHA1}$* and *fadD17$_{RHA1}$* were upregulated during growth of strain RHA1 on cholesterol, these CoA ligases are most likely involved in cholesterol side chain degradation. The cumulative inactivation of *fadD17$_{DSM43269}$* and *fadD19$_{DSM43269}$* in strain RG32, however, did not result in blocked cholesterol transformation. The most likely explanation is that the presence of multiple CoA ligase enzymes in strain RG32 accounts for the lack of phenotype of the *fadD17*/*fadD19* double mutant with cholesterol. It is generally known that CoA ligases are promiscuous enzymes, able to catalyze a wide variety of substrates with varying chain-lengths. The genome of strain RHA1 encodes an astonishing 73 putative CoA ligases, of which two (i.e. Ro05822 and Ro04675) have high sequence similarities with FadD17$_{RHA1}$ and FadD19$_{RHA1}$, respectively, and may be able to activate 5-cholestene-26-oic acid-3β-ol. Both Ro05822 and Ro04675 are highly conserved in all other available *Rhodococcus* genomes (as best reciprocal hits), and thus may also be present in *R. rhodochrous* DSM43269. Further investigations are needed to assess which CoA ligase enzymes are responsible for the first step of β-oxidation of the cholesterol side chain.

THE *LTP3* AND *LTP4* GENES ARE ESSENTIAL FOR DEGRADATION OF C24-BRANCHED CHAIN STEROLS

The *ltp3* gene was upregulated during growth of strain RHA1 on cholesterol (Chapter 2). However, deletion of a large set of β-oxidation genes, including *ltp3* and adjacent gene *ltp4*, in strain RHA1 (mutant strain MW1) did not affect growth on sterols (Chapter 4). Cloning of the *ltp3* and *ltp4* homologs from *R. rhodochrous* DSM43269 and subsequent deletion of these genes in derived strain RG32 revealed that mutants RG32Δ*ltp3* Δ*ltp4*, RG32Δ*ltp3* and RG32Δ*ltp4* are blocked in degradation of C24-branched sterol side chains, but still degrade the side chain of cholesterol (Chapter 5). Mutants RG32Δ*ltp3* and RG32Δ*ltp4* were successfully complemented by providing the respective deleted gene *in trans*, thus confirming that the observed phenotypes were due to deletion of the single genes only. Amino acid sequence comparisons have shown that Ltp3 and Ltp4, both annotated as non-specific lipid carrier proteins belonging to the family of thiolases, have relatively low similarities with the thiolase domain SCP-x of SCP-2. Two signature motifs, which include the catalytic residues (PF00108 and PF02803 (Finn *et al*., 2008) characteristic of thiolases, however, were not conserved in Ltp3 or Ltp4. Degradation of the phytosterol C24-branched chains does not occur by thiolytic C-C bond cleavage since the formation of a ketoacyl-CoA intermediate at C24, required for thiolytic activity, is not possible due to the maximal valence of the C24 atom. On the other hand, aldol lyase enzymes are capable of cleaving C-C

bonds of tertiary β-hydroxy CoA esters and thus C24-branched side chain degradation was proposed to occur by an aldol lyase reaction (Fujimoto *et al*., 1982a). We propose that *ltp3* and *ltp4* encode the latter enzymatic activity. It remains unclear whether Ltp3 and Ltp4 are involved in the same reaction, or that they have different functions. Another example of an aldol lyase enzyme is citrate lyase of *E. coli*; the active enzyme was shown to consist of three domains: CitE, CitF and CitG (Nilekani and SivaRaman, 1983). Similarly, Ltp3 and Ltp4 (and possibly other protein(s)) may form an aldol lyase enzyme complex that catalyzes C-C bond cleavage of C24-branched sterol side chains. Alternatively, one of the proteins may have aldol lyase activity and the other may be responsible for another essential step in C24-branched sterol side chain degradation. However, this cannot be concluded from the work presented in this thesis. The roles of Ltp3 and Ltp4 may be further studied by biochemical analysis of these proteins using substrates analogous to sterol side chain degradation intermediates.

FadA5 has been previously shown to be a thiolase responsible for cholesterol side chain degradation into AD and ADD in *M. tuberculosis* H37Rv. To substantiate the role of *fadA5* in *R. rhodochrous* DSM43269, we inactivated the $fadA5_{DSM43269}$ homolog in strain RG32 and in mutant strain RG32Δ*ltp3*Δ*ltp4*. Inactivation of $fadA5_{DSM43269}$ in both strains did not result in impaired cholesterol side chain degradation. Bioinformatic analysis revealed that the strain RHA1 genome encodes a close homolog of *fadA5*, i.e. *ro05815*, which is conserved in all available *Rhodococcus* genomes, but has no counterpart in strain H37Rv. Most likely, this homolog is also encoded by strain DSM43269. The presence of multiple thiolytic activities involved in cholesterol side chain degradation most likely accounts for the lack of phenotype observed for the $fadA5_{DSM43269}$ inactivated mutants.

FUTURE PROSPECTS

R. rhodochrous strain RG32, devoid of all KSH activity and unable to degrade steroid nuclei, was developed in our laboratory (Petrusma *et al*., in preparation) and shown to be capable of selective sterol side chain degradation. The strain RG32 phenotype is a suitable tool to study the role of genes in sterol side chain degradation specifically (as described in Chapters 3, 4 and 5). To identify the side chain degradation activities that still remain elusive, strain RG32 and derived mutants could be employed to characterize additional gene homologs from the strain RHA1 side chain degradation locus, or genes with high similarity to genes from this locus (such as the CoA ligases encoded by *ro05822* and *ro04675* (Chapter 4) and the thiolase encoded by *ro05815* (Chapter 5)). Likely, the cumulative inactivation of multiple genes encoding CoA ligases, acyl-CoA dehydrogenases or thiolases will result in strains blocked in cholesterol side chain degradation, and thus the identification of additional genes involved in this process. In view of the work presented in this thesis, it would be very interesting to know why deletion of single or multiple (β-oxidation) genes, e.g. $fadE26_{DSM43269}$, $fadE27_{DSM43269}$ and $ro04690_{DSM43269}$ (Chapter 4) or $fadA5_{DSM43269}$ (Chapter 5) in strain RG32 did not abolish cholesterol side chain degradation. Their strain RHA1 gene homologs are all located within the sterol side chain degradation locus and were generally

highly upregulated during growth of this strain on cholesterol, suggesting that they are involved in sterol catabolism. Amino acid sequence comparisons of the encoded proteins showed that most of the targeted genes have homologs encoding proteins with relatively high amino acid similarities within the strain RHA1 genome (e.g. Fig. 5, [Chapter 4] and Fig. 4, [Chapter 5]), which may be able to catalyze the same reaction. Therefore, targeted inactivation of such homologs in the aforementioned RG32 mutants may result in abolishment of cholesterol side chain degradation. In such a way, the genes and enzymes involved in cholesterol side chain degradation may finally be identified. Other enzyme activities involved in β-oxidation that need further exploration are enoyl-CoA hydratase and acyl-CoA/β-hydroxyacyl-CoA dehydrogenase. Enoyl-CoA hydratase activity is possibly encoded by *echA19* and/or *hsd4B* (*ro04688* and *ro04531* in strain RHA1, respectively). A multitude of genes in the genome of strain RHA1 are annotated as putative acyl-CoA/β-hydroxyacyl-CoA dehydrogenases. The sterol side chain degradation locus comprises many putative (acyl-CoA) dehydrogenases and oxidoreductases, including *ro04682*, *ro04686*, *ro04690*, *fadE27* (= *ro04692*), *fadE26* (= *ro04693*) and *hsd4A*.

An enzymatic activity that is of special interest in phytosterol side chain degradation is the C28-carboxylase (Fig. 1B, reaction 4). It was previously shown that in *Mycobacterium* sp. NRRL B-3805 the rate of β-sitosterol side chain degradation was significantly enhanced when the reaction was supplemented with HCO_3^-, indicating that carboxylation is the rate limiting step in this process (Chen, 1985). Therefore, addition of hydrogen carbonate in combination with over-expression of the enzyme system responsible for C24-branched sterol side chain degradation likely leads to enhanced conversion rates of phytosterols. Genetic information of such a carboxylase system is currently not available. An apparent operon encoding several putative carboxylase domains is located on plasmid pRHL2 comprising *ro10159-ro10162*. These genes were upregulated 2-6 fold during growth of strain RHA1 on cholesterol (Chapter 2) and are possible candidates to encode sterol C28-carboxylation activity in RHA1.

The mechanism and gene(s) responsible for the final step(s) of sterol side chain degradation, i.e. removal of the C3 side chain of 1,4-BNC, remains to be elucidated. However, once these activities and the encoding genes have been identified, mutants of *R. rhodochrous* strain RG32 can be constructed that may accumulate steroids of pharmaceutical interest, such as pregnanes. Steroids of the pregnane series (e.g. progesterone) are excellent precursors for corticosteroid synthesis. Conversely, these genes may be over-expressed in a production strain to achieve complete conversion of 1,4-BNC into ADD. It may also be that the incomplete conversion of 1,4-BNC into ADD is caused by product inhibition (at the enzyme level) or negative feedback of the genes encoding degradation of 1,4-BNC may occur. Thus, knowledge on the exact type of regulation of this process may also prove essential to successfully engineer sterol transforming *Actinobacteria*.

Besides the importance of bacterial sterol degradation in industrial processes, cholesterol metabolism is also relevant in view of the pathogenicity of *M. tuberculosis* (Chapter 2), the causative agent of tuberculosis (TB) in man, and possibly other pathogenic actinomycetes, such as *Rhodococcus equi* (van der Geize *et al*., 2008a). One third of the human population has a latent *M. tuberculosis* infection (WHO, 2005). In 10% of the cases the infection will progress into active disease, which is potentially fatal in immunocompromised individuals or if left untreated (Frieden *et al*., 2003; WHO, 2005). The widespread emergence of (multi and extensively) drug resistant strains is now considered a serious health threat, also in Western countries, and thus novel therapeutics are needed to combat these highly resistant strains. Several genes involved in *M. tuberculosis* sterol catabolism have recently been identified as pathogenicity factors, essential for the intracellular survival of the bacterium in macrophages. Since the enzymes encoded by these genes generally have no human counterparts, they may be suitable targets for novel therapeutics against strains resistant to the currently used antibiotics. Possible targets could be KshA (= Rv3526) or KstD (= Rv3537) enzymes, which were shown to be essential for long-term survival of *M. tuberculosis* in activated macrophages and in mice (Rengarajan *et al*., 2005; Hu *et al*., 2010).

The work presented in this thesis has resulted in identification of the first bacterial genes encoding enzymes catalyzing important steps in sterol side chain degradation and provides a clear contribution to our understanding of sterol degradation in *Actinobacteria*, including pathogenic strains such as *R. equi* and *M. tuberculosis*.

FINAL CONCLUSIONS

- A cholesterol catabolic gene cluster, consisting of 51 genes specifically expressed during growth on cholesterol was identified in the soil actinomycete *R. jostii* RHA1. These genes were found to be conserved in the obligatory pathogen *M. tuberculosis* H37Rv.

- Cytochrome P450 monooxygenase 125 (CYP125) is a steroid C26-hydroxylase essential for sterol C26-oic acid formation.

- FadD19 is a steroid-CoA ligase with an *in vivo* role in C24-branched sterol side chain degradation and able to activate 5-cholestene-26-oic acid-3β-ol and derivatives thereof *in vitro*.

- Our data suggest that in strain *R. rhodochrous* DSM43269 multiple CoA ligase isoenzymes are involved in cholesterol degradation.

- Ltp3 and Ltp4 are both essential for C24-branched sterol side chain degradation, possibly acting as aldol lyase(s). Their exact roles however remain to be verified.

- In contrast to *M. tuberculosis*, inactivation of *fadA5*, encoding a thiolase, in *R. rhodochrous* strain RG32 did not affect cholesterol side chain degradation, suggesting that multiple isoenzymes are present in this strain.

- Strain RG32 of *R. rhodochrous* DSM43269, with inactivated KSH activity and completely blocked in steroid ring degradation, is a useful tool to study the role of genes in sterol side chain degradation specifically.

- The large number of β-oxidation enzymes encoded by rhodococcal genomes likely have (partially) overlapping substrate specificities, since various multiple β-oxidation gene inactivation mutants in strain RG32 were not affected in (chole)sterol side chain degradation.

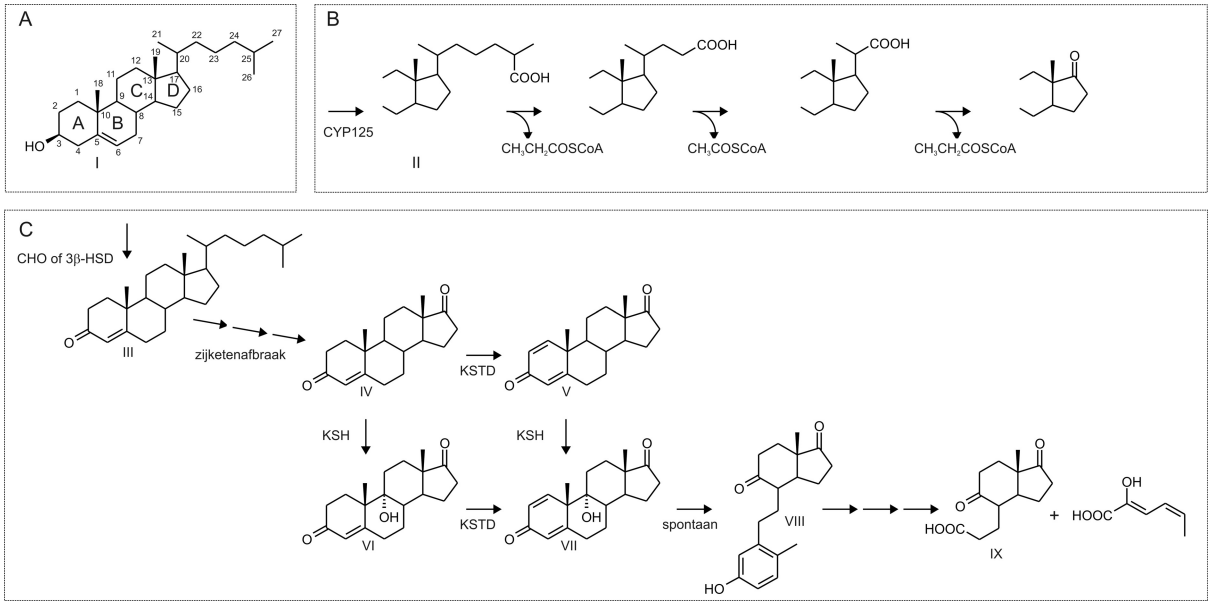
REFERENCES

References are listed on pages 119-140.

NEDERLANDSE SAMENVATTING

INTRODUCTIE

Vrijwel alle eukaryote organismen (planten, schimmels, dieren) bevatten sterolen in hun celmembranen die daar een belangrijke rol spelen in het bepalen van de structuur en doorlaatbaarheid. Sterolen zijn moleculen met een steroïde ringstructuur die bestaat uit 4 ringen, een hydroxylgroep op positie 3 en een alifatische zijketen op positie 17, en worden ook wel steroïde-alcoholen genoemd. Het bekendste sterol is cholesterol (Figuur 1A, molecuul I), dat o.a. in dieren voorkomt. Cholesterol en fytosterolen (uit planten) zijn de meest voorkomende voorbeelden van deze biomoleculen. Bacteriën en Archaea (prokaryoten) daarentegen bevatten doorgaans geen sterolen in hun celmembranen. Er zijn een aantal bacteriën, voornamelijk Actinobacteriën (waaronder *Mycobacterium*, *Nocardia* en *Rhodococcus*), die genetisch zodanig geëvolueerd zijn dat hun DNA codeert voor enzymen die sterolen afbreken. Hierdoor kunnen zij o.a. cholesterol als koolstof- en energiebron benutten (Hoofdstuk 1).

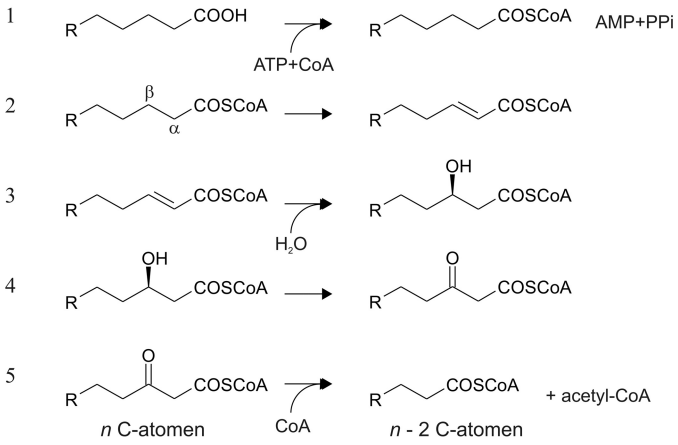


Figuur 1. Schematisch overzicht van de bacteriële afbraak van cholesterol. (A) Cholesterol bestaat uit de ringen A, B, C en D, met een hydroxyl (-OH) groep op positie 3 en een zijketen op positie 17. (B) Cholesterol zijketenafbraak begint met vorming van het C26-carbonzuur (molecuul II) door CYP125. Verdere afbraak vindt stapsgewijs plaats door middel van β-oxidatie, waarbij meerdere enzymen betrokken zijn. (C) Afbraak van de cholesterol ringstructuur begint met oxidatie en isomerisatie tot de 3-keto-Δ4 vorm (molecuul III) door cholesterol oxidase (CHO) of 3β-hydroxysteroïde dehydrogenase (3β-HSD). Uiteindelijk wordt ring B geopend door de werking van de enzymen 3-ketosteroïde 9α-hydroxylase (KSH) en 3-ketosteroïde Δ1-dehydrogenase (KSTD), waarbij 3-hydroxy-9,10-secoandrosta-1,3,5,(10)-triene-9,17-dione (3-HSA, molecuul VIII) gevormd wordt. 3-HSA wordt verder gemetaboliseerd tot 9,17-dioxo-1,2,3,4,10,19-hexanorandrostan-5-oic acid (DOHNAA, molecuul IX) en kan uiteindelijk compleet worden afgebroken tot H_2O en CO_2.

De afbraak van sterolen door bacteriën gebeurt door specifieke enzymen die de zijketen stapsgewijs verwijderen en speciale enzymen die de steroïde ringstructuur afbreken (zie Figuur 1). De volgorde waarin de enzymen van deze afbraakroutes het sterol molecuul aanvallen kan verschillen tussen verschillende bacteriën.

Vanuit een commercieel oogpunt is de bacteriële afbraak van sterolen van belang aangezien hierbij tussenproducten (steroïden) worden gevormd die als grondstof voor medicijnen (bijv. ontstekingsremmers) of anticonceptiva kunnen dienen. In de natuurlijke situatie zullen in de meeste bacteriën deze tussenproducten echter snel worden afgebroken tot moleculen die niet van farmaceutisch belang zijn. Voor de biotechnologische productie van steroïden uit sterolen is het essentieel dat de ringstructuur intact blijft en dat alleen de zijketen wordt afgebroken (= selectieve zijketenafbraak). Dit is mogelijk door gebruik te maken van gemuteerde bacteriële stammen waarin de ringafbraak geblokkeerd is of met behulp van chemische remmers die de enzymen die de ring afbreken blokkeren in hun activiteit. Vanwege de relevantie van sterol- en steroïdeafbraak voor de farmaceutische industrie is veel onderzoek gedaan naar de bacteriën, genen en enzymen die betrokken zijn bij deze processen.

Van o.a. *Rhodococcus* is bekend welke genen en enzymen verantwoordelijk zijn voor de afbraak van de steroïde ringstructuur. Welke genen betrokken zijn bij de afbraak van sterol zijketens was echter tot zeer recentelijk nog onbekend. Wel kon door de identificatie van de tussenproducten die gevormd worden en door het bestuderen van enzymreacties van sterol afbrekende cellen (of celextracten) het mechanisme van bacteriële sterol zijketenafbraak worden opgehelderd. Het proces waarmee dit gebeurt bleek analoog aan de β-oxidatie van vetzuren (Figuur 2).



Figuur 2. Reacties betrokken bij β-oxidatie: (1) ATP afhankelijke CoA activatie door CoA ligase. (2) Dehydrogenering van de α en β koolstofatomen door acyl-CoA dehydrogenase. (3) Hydrolyse van de dubbele binding door enoyl-CoA hydratase. (4) Tweede dehydrogeneringsreactie door β-hydroxyacyl-CoA dehydrogenase. (5) Thiolytische splitsing, door thiolase.

Het doel van het onderzoek beschreven in dit proefschrift was om gedetailleerde informatie te krijgen over de genen en enzymen die betrokken zijn bij bacteriële sterol zijketenafbraak. Zulke informatie is belangrijk vanuit fundamenteel oogpunt en voor het ontwikkelen van industrieel toepasbare steroïde productiestammen d.m.v. rationele cel “engineering”. Door

bijvoorbeeld de genen die coderen voor het zijketenafbraak proces tot overexpressie te brengen kan een verhoogde sterol transformatie efficiëntie worden behaald. Ook kunnen gericht een of meerdere van deze genen in een bacteriële stam uit worden geschakeld, wat zou kunnen leiden tot de vorming van nieuwe steroïde structuren die als grondstof voor nieuwe farmaceutica kunnen dienen (Hoofdstuk 1).
Aangezien het biologische mechanisme van bacteriële sterol zijketenafbraak bekend was (nl. β-oxidatie) zouden de gezochte genen zeer waarschijnlijk grote gelijkenis (homologie) vertonen met reeds bekende β-oxidatie genen van bijvoorbeeld vetzuur afbraak.
In mijn promotieonderzoek hebben we de bacteriën *Rhodococcus jostii* RHA1 en *Rhodococcus rhodochrous* DSM43269 bestudeerd, die beide zeer efficiënt sterolen en steroïden kunnen afbreken. De complete genetische informatie (genoomsequentie) van stam RHA1 was recentelijk opgehelderd. De RHA1 stam bleek echter zoveel β-oxidatie genen te bevatten dat het onbegonnen werk was om ze allemaal te onderzoeken. Echter, door gebruik te maken van transcriptoomanalyse was het mogelijk om te kijken welke genen “aan” of “uit” staan bij de groei van stam RHA1 op cholesterol. Met behulp van deze techniek werd een gencluster geïdentificeerd dat codeert voor de opname en afbraak van cholesterol (Hoofdstuk 2) in stam RHA1 (Figuur 3).

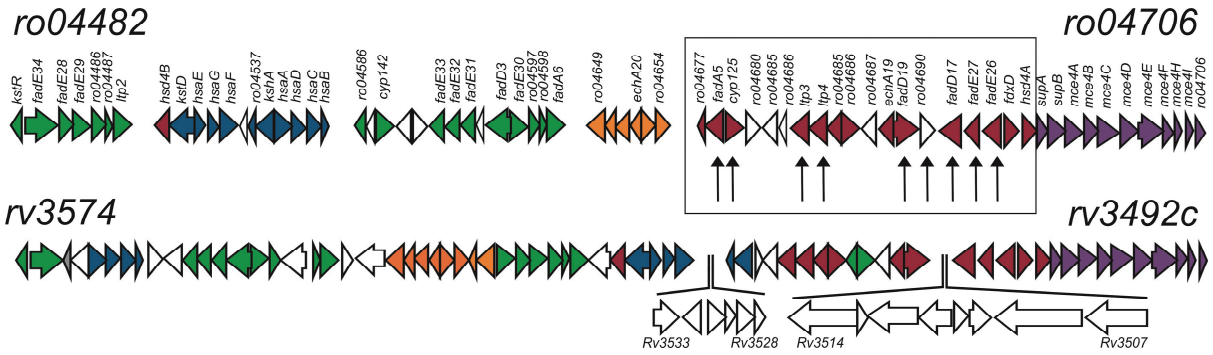
Van *R. rhodochrous* DSM43269 was bekend dat deze in de aanwezigheid van enzymremmers die het 3-ketosteroïde-9α-hydroxylase (KSH) enzym blokkeren de steroïden 1,4-androstadiene-3,17-dione (ADD, Figuur 4, molecuul V) en 3-oxo-23,24-bisnor-1,4-choladiene-22-carbonzuur (1,4-BNC, Figuur 4, molecuul IV) vormt uit sterolen. Stam DSM43269 is dus in staat tot selectieve sterol zijketenafbraak. Door de (vijf!) verschillende genen die coderen voor KSH activiteit in stam DSM43269 uit te schakelen werd mutant RG32 verkregen die ADD en 1,4-BNC ophoopt uit sterolen. De aldus gevormde steroïden worden vervolgens niet verder afgebroken door de RG32 stam. In Hoofdstukken 3, 4 en 5 staat beschreven hoe mutant RG32 succesvol werd gebruikt om gericht verschillende (β-oxidatie) genen te inactiveren om zo hun rol in sterol zijketenafbraak vast te stellen. Hierbij werden specifiek de (homologe) genen gekozen die in stam RHA1 aanstonden tijdens groei op cholesterol.

IDENTIFICATIE VAN EEN CHOLESTEROL AFBRAAK GENKLUSTER

Met behulp van DNA microarrays werd in *R. jostii* RHA1 een gencluster geïdentificeerd bestaande uit 51 genen met een hogere expressie bij groei op cholesterol (Figuur 3). Het gencluster, dat verantwoordelijk is voor cholesterol afbraak, bleek sterk geconserveerd in verwante bacteriën, inclusief de pathogene *Mycobacterium tuberculosis* H37Rv (Figuur 3). Verdere analyse van het cluster uit RHA1 toonde aan dat het genen bevat met een hoge gelijkenis aan de reeds bekende steroïde afbraak genen *kstD* en *kshA*. Ook kwamen zeven *hsa* genen versterkt tot expressie, deze genen coderen voor enzymen die betrokken zijn bij 3-HSA (Figuur 1C, molecuul VIII) afbraak in verwante Actinobacteriën. Sommige van de gecodeerde eiwitten waren geannoteerd als enzymen die betrokken zijn bij de afbraak van

de aromatische verbinding bifenyl, maar studies aan $HsaC_{H37Rv}$ en $HsaD_{H37Rv}$ toonden aan dat deze enzymen een veel hogere activiteit (~ 40 maal) hebben met steroïden dan met bifenyl metabolieten. Het cluster bevat daarnaast een operon dat verantwoordelijk is voor de opname van sterolen (van der Geize *et al*., 2007, [Hoofdstuk 2]; Mohn *et al*., 2008). Daarnaast werd een locus geïdentificeerd, bestaande uit de *ro04678-ro04695* genen die coderen voor β-oxidatie enzymen met een mogelijke rol in de afbraak van de zijketens van sterolen.

Het gencluster dat codeert voor cholesterol afbraak in de bodembacterie *R. jostii* RHA1 bleek sterk geconserveerd in de ziekteverwekkende bacterie *Mycobacterium tuberculosis*. Deze bacterie is de veroorzaker van tuberculose in de mens en veroorzaakt 1-2 miljoen doden per jaar. *M. tuberculosis* heeft als eigenschap dat deze jarenlang kan overleven in afweercellen (macrofagen) van de mens. Ondanks vele diepgaande studies was nog onduidelijk hoe deze pathogene bacterie zolang kan overleven in de vijandige omgeving van macrofagen. In deze studies waren genen geïdentificeerd die een belangrijke rol spelen bij de overleving in macrofagen, maar van veel genen was de fysiologische rol nog onbekend. Uit ons onderzoek bleek dat veel genen die belangrijk zijn voor de overleving van *M. tuberculosis* in macrofagen betrokken zijn bij cholesterol afbraak (Hoofdstuk 2). Latere studies door andere groepen hebben bevestigd dat *M. tuberculosis* cholesterol kan afbreken en dat deze eigenschap zeer belangrijk is voor pathogeniciteit (Brzostek *et al*., 2007, 2009; Chang *et al*., 2009; Hu *et al*., 2009; Lack *et al*., 2009; Nesbitt *et al*., 2010; Pandey and Sassetti, 2008; Yam *et al*., 2009). Aangezien de enzymen voor cholesterol afbraak in de mens niet voorkomen zijn dit mogelijke “targets” voor het ontwikkelen van nieuwe farmaceutica (enzymremmers) om de strijd aan te kunnen gaan met *M. tuberculosis* bacteriën die resistent zijn tegen de huidige gebruikte antibiotica.



Figuur 3. De cholesterol afbraak genclusters in *R. jostii* RHA1 (boven) en in *M. tuberculosis* H37Rv (onder). De 51 genen van stam RHA1 die verhoogd tot expressie komen bij groei op cholesterol liggen verspreid over een stuk DNA van 235 kb. In *M. tuberculosis* bestaat het cluster uit 82 aaneengesloten genen. De in dit proefschrift onderzochte RHA1 genen zijn in het kader met pijlen aangegeven.

CYP125 IS EEN STEROÏDE C26-HYDROXYLASE

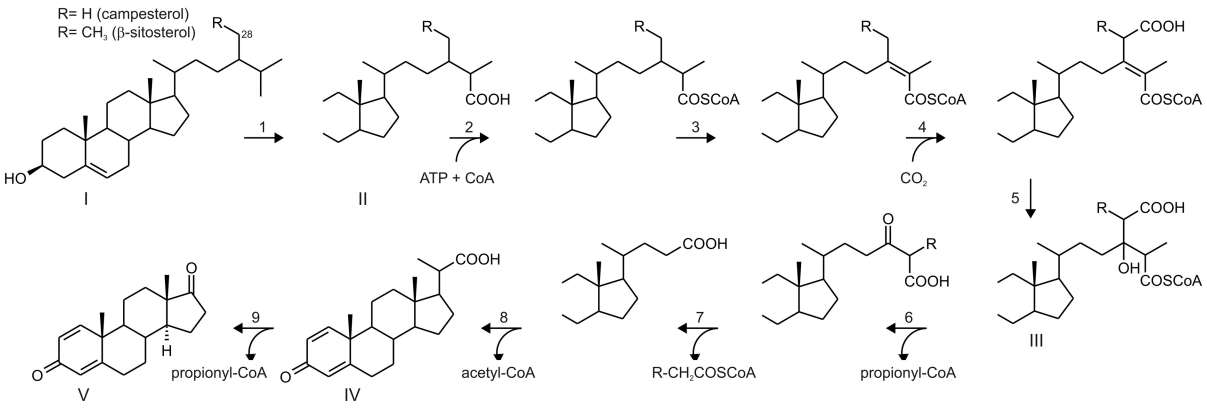
Het *cyp125* gen van stam RHA1 codeert voor een cytochroom P450 monooxygenase en was hoog opgereguleerd tijdens de groei op cholesterol (Hoofdstuk 2). Vanwege de ligging van *cyp125* naast β-oxidatie genen vermoedden wij dat dit gen betrokken was bij (chole)sterol zijketenafbraak, en wel als een sterol C26-hydroxylase enzym. Om dit te verifiëren werd de *cyp125* homoloog van *R. rhodochrous* DSM43269 geïnactiveerd in stam RG32. Gendisruptie mutant RG32Ω*cyp125* bleek inderdaad volledig geblokkeerd te zijn in zijketenafbraak van o.a. cholesterol en β-sitosterol (Hoofdstuk 3). Mutant RG32Ω*cyp125* bleek wel in staat om de zijketen af te breken van 5-cholestene-26-carbonzuur-3β-ol (Figuur 1B, molecuul II), een verwacht tussenproduct van cholesterol zijketenafbraak. Hieruit kon worden geconcludeerd dat CYP125 betrokken is bij de vorming van het C26-carbonzuur van sterolen, zeer waarschijnlijk als een C26-hydroxylase (Hoofdstuk 3). Ook werd van $CYP125_{RHA1}$ aangetoond dat het sterolen kon binden, maar hydroxylase activiteit met sterolen kon niet vastgesteld worden (Hoofdstuk 3). Latere studies met CYP125 uit *M. tuberculosis* H37Rv hebben bevestigd dat $CYP125_{H37Rv}$ in staat is om sterol zijketens te hydroxyleren op de C26 of C27 positie (Capyk *et al*., 2009; McLean *et al*., 2009) en deze verder te oxideren tot het carbonzuur (Ouellet *et al*., 2010; Johnston *et al*., 2010). CYP125 is dus verantwoordelijk voor de vorming van sterol C26-carbonzuren, wat essentieel is voor verdere afbraak van de zijketens d.m.v. β-oxidatie.

IDENTIFICATIE VAN $FADD19_{DSM43269}$ ALS EEN STEROÏDE-COA LIGASE

De eerste stap van β-oxidatie is CoA-activatie van het eindstandige carbonzuur en wordt gekatalyseerd door CoA ligase enzymen. Het sterol zijketenafbraak locus van stam RHA1 bevat twee genen die geannoteerd zijn als CoA ligase: *fadD17* en *fadD19* (Figuur 3). Beide genen waren opgereguleerd tijdens de groei op de cholesterol (Hoofdstuk 2). Om vast te stellen wat hun exacte rol is in sterol afbraak, werden de *fadD17* en *fadD19* genen van *R. rhodochrous* gekloneerd en geïnactiveerd in stam RG32. Opmerkelijk was dat mutanten waarin deze genen afzonderlijk van elkaar of tegelijkertijd waren geïnactiveerd niet geblokkeerd bleken te zijn in cholesterol zijketenafbraak. Mutant RG32Δ*fadD19* daarentegen bleek volledig geblokkeerd te zijn in de afbraak van de C24-vertakte zijketens van β-sitosterol en campesterol (Figuur 4). $FadD19_{DSM43269}$ heeft *in vivo* dus een essentiële rol in β-sitosterol en campesterol zijketenafbraak. Uit enzymassays bleek verder dat $FadD19_{DSM43269}$ actief is met 5-cholestene-26-carbonzuur-3β-ol (en de 3-keto-Δ4 vorm daarvan), waarmee werd bevestigd dat het enzym een steroïde-CoA ligase is. Helaas konden vertakte C24-sterol zijketens met een C26-carbonzuur niet getest worden, omdat deze niet (commercieel) beschikbaar zijn.

Aangezien $FadD19_{DSM43269}$ in staat is om de zijketens van 5-cholestene-26-carbonzuur-3β-ol metabolieten *in vitro* te katalyseren, en $fadD19_{RHA1}$ *in vivo* verhoogd tot expressie kwam tijdens groei op cholesterol, is FadD19 zeer waarschijnlijk wel betrokken bij cholesterol zijketenafbraak. Vanwege de ligging van $fadD17_{RHA1}$ in het zijketenafbraak locus en opregulatie van dit gen bij groei op cholesterol is het mogelijk dat FadD17 deze reactie ook

kan katalyseren. Onze resultaten suggereren daarom dat bij de zijketenafbraak van cholesterol de CoA activatie door meerdere CoA ligases uitgevoerd kan worden. Het is algemeen bekend dat CoA ligases in staat zijn om een breed scala van substraten (met bijvoorbeeld verschillende ketenlengtes) te activeren. Het genoom van stam RHA1 bevat 73 genen die geannoteerd zijn als CoA ligases. Hiervan hebben de eiwitten Ro05822 en Ro04675 een hoge aminozuursequentie gelijkenis met respectievelijk $FadD17_{RHA1}$ en $FadD19_{RHA1}$. Mogelijk zijn deze enzymen ook in staat om cholesterol C26-carbonzuren en daarmee de afbraak van de zijketen te activeren, bijvoorbeeld wanneer *fadD17* en/of *fadD19* geïnactiveerd zijn. Zowel Ro05822 als Ro04675 zijn sterk geconserveerd in alle andere beschikbare *Rhodococcus* genomen (als beste reciproke homologen) en zouden dus ook aanwezig kunnen zijn in *R. rhodochrous* DSM43269. Verder onderzoek is noodzakelijk om vast te kunnen stellen welke CoA ligases betrokken zijn bij cholesterol zijketenafbraak.



Figuur 4. Weergave van de afbraak van de C24-vertakte zijketens van β-sitosterol en campesterol. CYP125 is betrokken bij C26-carbonzuur vorming (stap 1). FadD19 is een steroïde-CoA ligase essentieel voor activatie van C24-vertakte zijketens (stap 2). Ltp3 en Ltp4 zijn essentieel voor afbraak van C24-vertakte sterol zijketens, waarschijnlijk bij stap 6.

LTP3 EN *LTP4* ZIJN ESSENTIEEL VOOR DE AFBRAAK VAN C24-VERTAKTE STEROLEN

Het *ltp3* gen, geannoteerd als aspecifieke lipide transporteiwit, behorende tot de familie van thiolases, kwam verhoogd tot expressie tijdens de groei van stam RHA1 op cholesterol (Hoofdstuk 2). Het uitschakelen van een groot aantal β-oxidatie genen, waaronder *ltp3* en het naastgelegen *ltp4* gen in stam RHA1 (mutant MW1) bleek niet van invloed op de groei op sterolen (Hoofdstuk 4). Het uitschakelen van *ltp3* en *ltp4* (geannoteerd als thiolase) in stam RG32 van *R. rhodochrous* bleek te leiden tot geblokkeerde afbraak van C24-vertakte zijketens (β-sitosterol, campesterol), maar niet van cholesterol (Hoofdstuk 5). Mutanten RG32Δ*ltp3* en RG32Δ*ltp4* werden gecomplementeerd door de wild type versies van de uitgeschakelde genen terug te brengen op een plasmide, waarmee werd bevestigd dat de waargenomen fenotypes veroorzaakt worden door het inactiveren van deze genen. Aminozuursequentie vergelijkingen wezen uit dat zowel Ltp3 en Ltp4 relatief lage overeenkomsten hebben met eerder gekarakteriseerde thiolases. Ook bleken twee

kenmerkende aminozuursequentie motieven (PF00108 en PF02803 (Finn *et al*., 2008)), die o.a. de katalytische residuen van thiolases bevatten, niet geconserveerd te zijn in Ltp3 en Ltp4. Daarnaast is in eerder onderzoek beschreven dat de bacteriële afbraak van C24-vertakte sterolzijketens niet door thiolytische splitsing plaatsvindt, maar door een aldol lyase reactie (Fujimoto *et al*., 1982a). Thiolytische splitsing is chemisch niet mogelijk aangezien de vorming van een 3-ketoacyl-CoA, noodzakelijk voor thiolytische activiteit, niet plaats kan hebben vanwege de vierwaardigheid van het C24-atoom (zie Figuur 4, molecuul III). Aldol lyase enzymen zijn in staat om C-C bindingen te splitsen van tertiaire β-hydroxyl CoA-esters en dus werd voor de afbraak van C24-vertakte sterol zijketens voorgesteld dat deze plaatsvindt d.m.v. een aldol lyase reactie (Figuur 4, reactie 6; Fujimoto *et al*., 1982a). Onze resultaten suggereren dat *ltp3* en/of *ltp4* coderen voor laatstgenoemde enzymatische activiteit. Biochemische karakterisatie van Ltp3 en Ltp4 is echter nodig om hun exacte rol te bepalen.

FadA5 is een thiolase die betrokken is bij de zijketenafbraak van cholesterol tot 4-androstene-3,17-dione (Figuur 1C, molecuul IV) in *M. tuberculosis* H37Rv (Nesbitt *et al*., 2009). Om de rol van *fadA5* in *R. rhodochrous* DSM43269 te verifiëren werd $fadA5_{DSM43269}$ geïnactiveerd in stam RG32 en in mutant RG32Δ*ltp3*Δ*ltp4*. Hieruit bleek dat inactivatie van $fadA5_{DSM43269}$ in beide gevallen in *R. rhodochrous* niet leidde tot geblokkeerde cholesterol zijketenafbraak. Uit bioinformatische analyse bleek dat het genoom van *R. jostii* RHA1 een homoloog van *fadA5* bevat, te weten *ro05815*. Dit gen is geconserveerd in alle beschikbare *Rhodococcus* genomen, maar is afwezig in *M. tuberculosis* H37Rv. Mogelijk is een *ro05815* homoloog ook aanwezig in stam DSM43269 en kan zijn aanwezigheid (en mogelijk dat van andere thiolases) het ontbreken van een fenotype voor onze $fadA5_{DSM43269}$ geïnactiveerde mutanten verklaren.

SLOTCONCLUSIES

Dit proefschrift beschrijft voor het eerst de identificatie en moleculaire karakterisatie van bacteriële genen die coderen voor enzymen die belangrijke stappen in sterol zijketenafbraak katalyseren en levert aldus een belangrijke bijdrage aan ons begrip van sterol afbraak in Actinobacteriën, met inbegrip van pathogene stammen als *Rhodococcus equi* en *M. tuberculosis*.

- Een cholesterol afbraak gen cluster, bestaande uit 51 genen die specifiek tot expressie kwamen tijdens de groei op cholesterol werd geïdentificeerd in de bodembacterie *R. jostii* RHA1. De gevonden genen bleken geconserveerd te zijn in de ziekteverwekkende bacterie *M. tuberculosis* H37Rv.

- Het cytochroom P450 monooxygenase 125 (CYP125) is een steroïde C26-hydroxylase met een essentiële rol in sterol C26-carbonzuur vorming en cholesterol zijketenafbraak.

- FadD19 is een steroïde-CoA ligase die *in vivo* essentieel is voor de afbraak van C24-vertakte sterol zijketens en die *in vitro* in staat is 5-cholestene-26-carbonzuur-3β-ol en afgeleiden daarvan te activeren.

- Ons onderzoek suggereert dat in stam *R. rhodochrous* DSM43269 meerdere CoA ligase isoenzymen betrokken zijn bij cholesterol zijketenafbraak.

- Ltp3 en Ltp4 hebben beide een essentiële rol in de afbraak van C24-vertakte sterol zijketens, waarschijnlijk als aldol lyase(s). Hun exacte functie is echter nog onduidelijk.

- In *R. rhodochrous* stam RG32, had inactivatie van *fadA5*, dat codeert voor een thiolase, geen effect op cholesterol zijketenafbraak, in tegenstelling tot *M. tuberculosis*. Waarschijnlijk zijn meerdere FadA5 isoenzymen aanwezig in *R. rhodochrous*.

- Stam RG32 van *R. rhodochrous* DSM43269, met geïnactiveerde KSH activiteit, is zeer geschikt om specifiek de rol van genen in sterol zijketenafbraak te bestuderen.

- *Rhodococcus* genomen coderen voor een zeer groot aantal β-oxidatie enzymen die zeer waarschijnlijk elkaar overlappen in hun substraat specificiteiten. Dit kan verklaren waarom mutanten van stam RG32 met meervoudige β-oxidatie gen inactivaties niet geblokkeerd zijn in (chole)sterol zijketenafbraak.

REFERENTIES

Referenties staan vermeld op pagina’s 119-140.

REFERENCES

Aaronson, S. (1964) A role for a sterol and a sterol precursor in the bacterium *Rhodopseudomonas palustris*. *J. Gen. Microbiol.* **37:** 225-229

Adams, J. (2002) Preclinical and clinical evaluation of proteasome inhibitor PS-341 for the treatment of cancer. *Curr. Opin. Chem. Biol.* **6:** 493-500

Adler, J.J. & Rose D.N. (1996) *Transmission and Pathogenesis of Tuberculosis.* In: Tuberculosis. (WN Rom, SM Garay, Eds.), Little, Brown, and Co. Boston

Ahmed, S., and Johri, B.N. (1993a) Microbial transformation of steroids in organic media. *Indian J. Chem.* (sect B) **32:** 67-69

Ahmed, S., Roy, P.K., and Basu, S.K. (1993b) Cholesterol side-chain cleavage by immobilized cells of *Rhodococcus equi* DSM 89-133. *Indian J. Exp. Biol.* **31:** 319-322

Aihara, H., Watanabe, K., and Nakamura, R. (1986) Characterization of production of cholesterol oxidases in three *Rhodococcus* strains. *J. Appl. Bacteriol.* **61:** 269-274

Akhtar, N., Ghauri, M.A., Anwar, M.A., and Akhtar, K. (2009) Analysis of the dibenzothiophene metabolic pathway in a newly isolated *Rhodococcus* spp. *FEMS Microbiol. Lett.* **301:** 95-102

Akihisa, T., Kokke, W., and Tamura, T. (1991) Naturally occurring sterols and related compounds from plants, in *Physiology and Biochemistry of Sterols* (Patterson, G.W. and Nes, W.D., eds), pp. 172-228, American Oil Chemists' Society

Alberts, B., Bray, D., Lewis, J., Raft, M., Roberts, K., and Watson, J.D. (1994) Molecular Biology of the Cell, 3rd Edition, Garland Publishing Inc., New York

Ambrus, G., Jekkel, A., Ilkoy, E., Horvath, G., and Bocksei, Z. (1995) Novel 26-oxygenated products in microbial degradation of ergosterol. *Steroids* **60:** 626-629

Anderton, M.C., Bhakta, S., Besra, G.S., Jeavons, P., Eltis, L.D., and Sim, E. (2006) Characterization of the putative operon containing arylamine N-acetyltransferase (nat) in *Mycobacterium bovis* BCG. *Mol. Microbiol.* **59:** 181-192

Antonenkov, V.D., van Veldhoven, P.P., Waelkens, E., and Mannaerts, G.P. (1997) Substrate specificities of 3-oxoacyl-CoA thiolase A and sterol carrier protein 2/3-oxoacyl-CoA thiolase purified from normal rat liver peroxisomes. *J. Biol. Chem.* **272:** 26023-26031

Arima, K., Nakamatsu, T., and Beppu, T. (1978) Microbial production of 3-oxobisnorchola-1,4-dien-22-oic acid. *Agric. Biol. Chem.* **42:** 411-416

Arima, K., Nagasawa, H., Bae, H., and Tamura, G. (1969) Microbial transformation of sterols. I. Decomposition of cholesterol by microorganisms. *Agr. Biol. Chem.* **33:** 1636-1643

Arora, P., Vats,A., Saxena, P., Mohanty, D., and Gokhale, R.S. (2005) Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis. *J. Am. Chem. Soc.* **127:** 9388-9389

Arruda, S., Bomfim, G., Knights, R., Huima-Byron, T., and Riley, L.W. (1993) Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. *Science* **261:** 1454-1457

Atrat, P.G., Koch, B., Szekalla, B., and Hoerhold, S.C. (1992) Application of newly synthesized detergents in the side-chain degradation of plant sterols by *Mycobacterium fortuitum*. *J. Basic Microbiol.* **132:** 147-157

Av-Gay, Y., and Sobouti, R. (2000) Cholesterol is accumulated by mycobacteria but its degradation is limited to non-pathogenic fast-growing mycobacteria. *Can. J. Microbiol*. **46:** 826-831

Banerjee, A., Sharma, R., and Banerjee, U.C. (2002) The nitrile-degrading enzymes: current status and future prospects. *Appl. Microbiol. Biotechnol.* **60:** 33-44

Bardone, M.R., Paternoster, M., and Coronelli, C. (1978) Teichomycins, new antibiotics from *Actinoplanes teicomyceticus* nov. sp. *J. Antibiot.* **31:** 170-177

Bell, K.S., Philp, J.C., Aw, D.W., and Christofi, N. (1998) The genus *Rhodococcus*. *J. Appl. Microbiol.* **85:** 195-210

Bendtsen, J.D., Nielsen, H., Von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340:** 783-795

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002) Biochemistry, fifth ed. pp. 731-738 W.H. Freeman and Company, NY

Bicca, F.C., Fleck, L.C., and Ayub, M.A.Z. (1999) Production of biosurfactant by hydrocarbon degrading *Rhodococcus ruber* and *Rhodococcus erythropolis*. *Rev. Microbiol.* **30:** 231-236

Biellmann, J.F. (2001) Resolution of alcohols by cholesterol oxidase from *Rhodococcus erythropolis*: lack of enantiospecificity for the steroids. *Chirality* **13:** 34-39

Birke, M., Schildbach, A., Heller, I., Hoerhold, C., Seidel, L., and Atart, P. (1992) Microbial manufacture and recovery of 9alpha-hydroxyandrost-4-ene-3,17-dione. (East) German Patent 298, 279 (Jena Pharm. GmbH., Jena)

Birkenmaier, A., Holert, J., Erdbrink, H., Moeller, H.M., Friemel, A., Schoenenberger, R., Suter, M.J., Klebensberger, J., and Philipp, B. (2007) Biochemical and genetic investigation of initial reactions in

aerobic degradation of the bile acid cholate in *Pseudomonas* sp. strain Chol1. *J. Bacteriol.* **189:** 7165-7173

Bode, H.B., and Müller, R. (2005) The impact of bacterial genomics on natural product research. *Angew. Chem. Int. Ed. Engl.* **44:** 6828-6846

Bode, H. B., Zeggel, B. Silakowski, B., Wenzel, S. C. , Reichenbach, H., and Müller, R. (2003) Steroid biosynthesis in prokaryotes: identification of myxobacterial steroids and cloning of the first bacterial 2,3(S)-oxidosqualene cyclase from the myxobacterium *Stigmatella aurantiaca*. *Mol. Microbiol.* **47:** 471-481

Borman, E.A., Redikulifsev, Yu V., Koshcheenko, K.A., Turuta, A.M., and Kamernitskii, A.V. (1992) Transformation of sitosterol to 9alpha-hydroxy-androstene-dione by *Mycobacterium* sp. 207 cells in the presence of an adsorption resin. *Prikl. Biokhim. Mikrobiol.* **28:** 551-556

Brands, S.J. (comp.) (1989-2005) Systema Naturae 2000. Amsterdam, The Netherlands. Available at: http://sn2000.taxonomy.nl/

Brooks, C.J.W. (1979) Some aspects of mass spectrometry in research on steroids. *Phil. Trans. R. Soc. Lond. A* **293:** 53-67

Bun-ya, M., Maebuchi, M., Kamiryo, T., Kurosawa, T., Sato, M., Tohma, M., Jiang, L.L., and Hashimoto, T. (1998) Thiolase involved in bile acid formation. *J. Biochem.* **123:** 347-352

Brzostek, A., Dziadek, B., Rumijowska-Galewicz, A., Pawelczyk, J., and Dziadek, J. (2007) Cholesterol oxidase is required for virulence of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **275:** 106-112

Brzostek, A., Pawelczyk, J., Rumijowska-Galewicz, A., Dziadek, B., and Dziadek, J. (2009) *Mycobacterium tuberculosis* is able to accumulate and utilize cholesterol. *J. Bacteriol.* **191:** 6584-6591

Brzostek, A., Sliwiński, T., Rumijowska-Galewicz, A., Korycka-Machała, M., and Dziadek, J. (2005) Identification and targeted disruption of the gene encoding the main 3-ketosteroid dehydrogenase in *Mycobacterium smegmatis*. *Microbiology* **151:** 2393-2402

Buckland, B.C., Drew, S.W., Connors, N.C., Chartrain, M.M., Lee, C., Salmon, P.M., Gbewonyo, K., Zhou, W., Gailliot, P., Singhvi, R., Olewinski, R.C. Jr., Sun, W.J., Reddy J., Zhang, J., Jackey, B.A., Taylor, C., Goklen, K.E., Junker, B., and Greasham, R.L. (1999) Microbial conversion of indene to indandiol: a key intermediate in the synthesis of CRIXIVAN. *Metab. Eng.* **1:** 63-74

Camus, J.C., Pryor, M.J., Medigue, C., and Cole, S.T. (2002) Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148:** 2967-2973

Capyk, J.K., D'Angelo, I., Strynadka, N.C., and Eltis, L.D. (2009) Characterization of 3-ketosteroid 9alpha-hydroxylase, a Rieske oxygenase in the cholesterol degradation pathway of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **284:** 9937-9946

Capyk, J.K., Kalscheuer, R., Stewart, G.R., Liu, J., Kwon, H., Zhao, R., Okamoto, S., Jacobs, W.R. Jr., Eltis, L.D., and Mohn, W.W. (2009) Mycobacterial cytochrome P450 125 (CYP125) catalyzes the terminal hydroxylation of C27 steroids. *J. Biol. Chem.* **284:** 35534-35542

Challis, G. (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* **154:** 1555-1569

Chang, J.C., Harik, N.S., Liao, R.P., and Sherman, D.R. (2007) Identification of Mycobacterial genes that alter growth and pathology in macrophages and in mice. *J. Infect. Dis.* **196:** 788-795

Chang, J.C., Miner, M.D., Pandey, A.K., Gill, W.P., Harik, N.S., Sassetti, C.M., and Sherman, D.R. (2009) *igr* genes and *Mycobacterium tuberculosis* cholesterol metabolism. *J. Bacteriol.* **191:** 5232-5239

Chang, F.N., and Sih, C.J. (1964) Mechanisms of steroid oxidation by microorganisms. 7. Properties of the 9alpha-hydroxylase. *Biochemistry* **3:** 1551-1557

Chang, K.H., Xiang, H., and Dunaway-Mariano, D. (1997) Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with the *Pseudomonas* sp. strain CBS3 4-chlorobenzoate:coenzyme A ligase. *Biochemistry* **36:** 15650-15659

Chauhan, D., Hideshima, T., and Anderson, K.C. (2006) A novel proteasome inhibitor NPI-0052 as an anticancer therapy. *Br. J. Cancer* **95:** 61-65

Chen, C.S. (1985) The mechanism of degradation of side chains of phytosterols by microorganisms. Ph.D. thesis. University of Wisconsin-Madison, Ann Arbor, M.I.

Chen, Y.-R., Huan, H.-H., Cheng, T.-F., Tang, T.-Y., and Liu, W.-H. (2006) Expression of a cholesterol oxidase gene from *Arthrobacter simplex* in *Escherichia coli* and *Pichia pastoris*. *Enzyme Microb. Technol.* **39:** 258-262

Chiang, Y.R., Ismail, W., Heintz, D., Schaeffer, C., van Dorsselaer, A., and Fuchs, G. (2008) Study of anoxic and oxic cholesterol metabolism by *Sterolibacterium denitrificans*. *J. Bacteriol.* **190:** 905-914

Chitale, S., Ehrt, S., Kawamura, I., Fujimura, T., Shimono, N., Anand, N., Lu, S., Cohen-Gould, L., and Riley, L.W. (2001) Recombinant *Mycobacterium tuberculosis* protein associated with mammalian cell entry. *Cell. Microbiol.* **3:** 247-254

Choi, K.Y., Kim, D., Sul, W.J., Chae, J.C., Zylstra, G.J., Kim, Y.M., and Kim E. (2005) Molecular and biochemical analysis of phthalate and terephthalate degradation by *Rhodococcus* sp. strain DK17. *FEMS Microbiol. Lett.* **252:** 207-213

Clark-Curtiss, J.E., and Haydel, S.E. (2003) Molecular genetics of *Mycobacterium tuberculosis* pathogenesis. *Annu. Rev. Microbiol.* **57:** 517-549

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E. 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., and Barrell, B.G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393:** 537-544

Crespi, M., Messens, E., Caplan, A.B., van Montagu, M., and Desomer, J. (1992) Fasciation induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. *EMBO J.* **11:** 795-804

Croteau, N., and Vrielink, A. (1996) Crystallization and preliminary X-ray analysis of cholesterol oxidase from *Brevibacterium sterolicum* containing covalently bound FAD. *J. Struct. Biol.* **116:** 317-319

Dabbs, E.R. (1998) Cloning of genes that have environmental and clinical importance from rhodococci and related bacteria. *Antonie van Leeuwenhoek* **74:** 155-167

Daffe M, McNeil M, Brennan PJ. (1993) Major structural features of the cell wall arabinogalactans of *Mycobacterium*, *Rhodococcus*, and *Nocardia* spp. *Carbohydr. Res.* **249:** 383-398

Daubner, S.C., Gadda, G. Valley, M.P., and Fitzpatrick, P.F. (2002) Cloning of nitroalkane oxidase from *Fusarium oxysporum* identifies a new member of the acyl-CoA dehydrogenase superfamily. *Proc. Natl. Acad. Sci. USA* **99:** 2702-2707

de Carvalho, C.C., and da Fonseca, M.M. (2005) The remarkable *Rhodococcus erythropolis*. *Appl. Microbiol. Biotechnol.* **67:** 715-726

de Chastellier, C., and Thilo, L. (2006) Cholesterol depletion in *Mycobacterium avium*-infected macrophages overcomes the block in phagosome maturation and leads to the reversible sequestration of viable mycobacteria in phagolysosome-derived autophagic vacuoles. *Cell. Microbiol.* **8:** 242-256

Dhar, A., and Samanta, T.B. (1993) Novel oxidative cleavage of C17-C20 bond in pregnane by a *Pseudomonas* sp. *J. Steroid. Biochem. Mol. Biol.* **44:** 101-104

Dias, A.C.P., Fernandes, P., Cabral, J.M.S., and Pinheiro, H.M. (2002) Isolation of a biodegradable sterol-rich fraction from industrial wastes. *Bioresour. Technol.* **82:** 253-260

Do, R., Kiss, R.S., Gaudet, D., and Engert, J.C. (2009) Squalene synthase: a critical enzyme in the cholesterol biosynthesis pathway. *Clin. Genet.* **75:** 19-29

Dodson, R. M., and Muir, R. D. (1958a) Microbiological transformations. II. The microbiological aromatization of steroids. *J. Am. Chem. Soc.* **80:** 5004-5005

Dodson R.M., and Muir, R.D. (1958b) Microbiological transformations. III. The hydroxylation of steroids at C-9. *J. Am. Chem. Soc.* **80:** 6148

Dovbnya, D.V., Egorova, O.V., and Donova, M.V. (2010) Microbial side-chain degradation of ergosterol and its 3-substituted derivatives: a new route for obtaining of deltanoids. *Steroids* **75:** 653-658

Dresen, C., Lin, L.Y., D'Angelo, I., Tocheva, E.I., Strynadka, N., and Eltis, L.D. (2010) A flavin-dependent monooxygenase from *Mycobacterium tuberculosis* involved in cholesterol catabolism. *J. Biol. Chem.* **285:** 22264-22275

Dutta, R.K., Roy, M., and Singh, H.D. (1992) Metabolic blocks in the degradation of beta-sitosterol by a plasmid-cured strain of *Arthrobacter oxydans. J. Basic Microbiol.* **32:** 167-176

Engel, R., and Knorr, D. (2004) Production of liquid, water-dispersible, phytosterol formulations for increased dose response in food systems. *Eng. Life. Sci.* **4:** 374-377

Feling, R.H., Buchanan, G.O., Mincer, T.J., Kauffman, C.A., Jensen, P.R., and Fenical, W. (2003) Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinospora*. *Angew. Chem. Int. Ed. Engl.* **42:** 355-357

Fernandes, P., Cruz, A., Angelova, B., Pinheiro, H. M., and Cabral, J. M. S. (2003) Microbial conversion of steroid compounds: recent developments. *Enzyme Microb. Technol.* **32:** 688-705

Fernandes, P.J., Powell, J.A., and Archer J.A. (2001) Construction of *Rhodococcus* random mutagenesis libraries using Tn5 transposition complexes. *Microbiology* **147:** 2529-2536

Fernández de Las Heras, L., Mascaraque, V., García Fernández, E., Navarro-Llorens, J.M., Perera, J., and Drzyzga, O. (2010) ChoG is the main inducible extracellular cholesterol oxidase of *Rhodococcus* sp. strain CECT3014. *Microbiol. Res.* In press

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., and Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res.* **36:** D281-288

Finnerty, W.R. (1992) The biology and genetics of the genus *Rhodococcus*. *Annu. Rev. Microbiol.* **46:** 193-218

Flynn, J.L., and Chan, J. (2001) Tuberculosis: latency and reactivation. *Infect. Immun.* **69:** 4195-4201

Fortin, P.D., Lo, A.T.F., Haro, M.A., Kaschabek, S.R., Reineke, W., and Eltis, L.D. (2005) Evolutionarily divergent extradiol dioxygenases possess higher specificities for polychlorinated biphenyl metabolites. *J. Bacteriol*. **187:** 415-421

Frieden, T.R., Sterling, T.R., Munsiff, S.S., Watt, C.J., and Dye, C. (2003) Tuberculosis. *Lancet* **362:** 887-899

Fujimoto, Y., Chen, C.S., Szeleczky, Z., Di Tullio, D., and Sih, C.J. (1982a) Microbial degradation of the phytosterol side chain. I. Enzymatic conversion of 3-oxo-24-ethylcholest-4-en-26-oic acid into 3-oxochol-4-en-24-oic acid and androst-4-ene-3,17-dione. *J. Am. Chem. Soc.* **104:** 4718-4720

Fujimoto, Y., Chen, C.S., Gopalan, A.S., and Sih, C.J. (1982b) Microbial degradation of the phytosterol side chain. II. Incorporation of $NaH^{14}CO_3$ onto the C-28 position. *J. Am. Chem. Soc.* **104:** 4720-4722

Fujishiro, K., Ohta, T., Hasegawa, M., Yamaguchi, K., Mizukami, T., Uwajima, T., and Ota, T. (1990) Isolation and identification of the gene of cholesterol oxidase from *Brevibacterium sterolicum* ATCC 21387, a widely used enzyme in clinical analysis. *Biochem. Biophys. Res. Commun.* **172:** 721-727

Gatfield, J., and Pieters, J. (2000) Essential role for cholesterol in entry of mycobacteria into macrophages. *Science* **288:** 1647-1650

Gendron, J.M., and Wang, Z.Y. (2007) Multiple mechanisms modulate brassinosteroid signaling, *Curr. Opin. Plant Biol.* **10:** 436-441

Geraci, J.E., Heilman, F.R., Nichols, D.R., Ross, G.T., and Wellman, W.E. (1956) Some laboratory and clinical experiences with a new antibiotic, vancomycin. *Mayo Clin. Proc.* **31:** 564-582

Gibson, D.T., Wang, K.C., Sih, C.J., and Whitlock, H.Jr. (1966) Mechanisms of steroid oxidation by microorganisms. IX. On the mechanism of ring A cleavage in the degradation of 9,10-seco steroids by microorganisms. *J. Biol. Chem.* 241: 551-559

Gonçalves, E.R., Hara, H., Miyazawa, D., Davies, J.E., Eltis, L.D., and Mohn, W.W. (2006) Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol.* **72:** 6183-6193

Goodfellow, M., Alderson, G., and Chun, J. (1998) Rhodococcal systematics: problems and developments. *Antonie Van Leeuwenhoek* **74:** 3-20

Gottelt, M., Kol, S., Gomez-Escribano, J.P., Bibb, M., and Takano, E. (2010) Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology* **156:** 2343-2353

Gottschaldt, B., Grosse, H.H., Hoerhold, C., Welzker, M., Naumann, H., Heller, I., Birke, M., and Plonka, G. (1993) German Patent 301, 740 (Institut für Microbiologie und experimentelle Therapie; Jenapharm GmbH JENA)

Groot, P.H., Scholte, H.R., and Hülsmann, W.C. (1976) Fatty acid activation: specificity, localization, and function. *Adv. Lipid Res.* **14:** 75-126

Gürtler, V., Mayall, B.C., and Seviour, R. (2004) Can whole genome analysis refine the taxonomy of the genus *Rhodococcus*? *FEMS Microbiol. Rev.* **28:** 377-403

Haapalainen, A.M., Meriläinen, G., and Wierenga, R.K. (2006) The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends Biochem. Sci.* **31:** 64-71

Hacking, A. J., and Quayle, J.R. (1974) Purification and properties of malyl-coenzyme A lyase from *Pseudomonas* AM1. *Biochem. J.* **139:** 399-405

Helliwell, C.A., Poole, A., Peacock, W.J., and Dennis, E.S. (1999) *Arabidopsis ent*-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant. Physiol.* **119:** 507-510

Helliwell, C.A., Chandler, P.M., Poole, A., Dennis, E.S., and Peacock, W.J. (2001) The CYP88A cytochrome P450, entkaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc. Natl. Acad. Sci. USA* **98:** 2065-2070

Honeywell, E.M., and Bills, C.E. (1932) Cerevisterol, a sterol accompanying ergosterol in yeast. *J. Biol. Chem.* **99:** 71-78

Hopwood, D.A. (2007) *Streptomyces* in nature and medicine: the antibiotic makers. Oxford University Press, New York, NY

Horinouchi, M., Hayashi, T., Yamamoto, T., and Kudo, T. (2003) A new bacterial steroid degradation gene cluster in *Comamonas testosteroni* TA441 which consists of aromatic-compound degradation genes for seco-steroids and 3-ketosteroid dehydrogenase genes. *Appl. Environ. Microbiol.* **69:** 4421-4430

Horinouchi, S., Ishizuka, H., and Beppu, T. (1991) Cloning, nucleotide sequence, and transcriptional analysis of the NAD(P)-dependent cholesterol dehydrogenase gene from a *Nocardia* sp. and its hyperexpression in *Streptomyces* spp. *Appl. Environ. Microbiol.* **57:** 1386-1393

Horinouchi, M., Kurita, T., Yamamoto, T., Hatori, E., Hayashi, T., and Kudo, T. (2004) Steroid degradation gene cluster of *Comamonas testosteroni* consisting of 18 putative genes from meta-cleavage enzyme gene *tesB* to regulator gene *tesR*. *Biochem Biophys. Res. Commun.* **324:** 597-604

Horinouchi, M., Yamamoto, T., Taguchi, K., Arai, H., and Kudo, T. (2001) Meta-cleavage enzyme gene *tesB* is necessary for testosterone degradation in *Comamonas testosteroni* TA441. *Microbiology* **147:** 3367-3375

Hrycay, E.G., Gustafsson, J.A., Ingelman-Sundberg, M., Ernster, L. (1975) Sodium periodate, sodium chloride, organic hydroperoxides, and H_2O_2 as hydroxylating agents in steroid hydroxylation

reactions catalyzed by partially purified cytochrome P-450. *Biochem. Biophys. Res. Commun.* **66:** 209-216

Hu, Y., van der Geize, R., Besra, G.S., Gurcha, S.S, Liu, A., Rohde, M., Singh, M., and Coates, A. (2009) 3-Ketosteroid 9alpha-hydroxylase is an essential factor in the pathogenesis of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **75:** 107-121

Ishikawa, J., Yamashita, A., Mikami, Y., Hoshino, Y., Kurita, H., Hotta, K., Shiba, T., and Hattori, M. (2004) The complete genomic sequence of *Nocardia farcinica* IFM 10152. *Proc. Natl. Acad. Sci. USA* **101:** 14925-14930

Ishizaki, T., Hirayama, N., Shinkawa, H., Nimi, O., and Murooka, Y. (1989) Nucleotide sequence of the gene for cholesterol oxidase from a *Streptomyces* sp. *J. Bacteriol.* **171:** 596-601

Itagaki, E., Hatta, T., Wakabayashi, T., and Suzuki, K. (1990a) Spectral properties of 3-ketosteroid delta1-dehydrogenase from *Nocardia corallina*. *Biochim. Biophys. Acta* **1040:** 281-286

Itagaki, E., Wakabayashi, T., and Hatta, T. (1990b) Purification and characterization of 3-ketosteroid delta1-dehydrogenase from *Nocardia corallina*. *Biochim. Biophys. Acta* **1038:** 60-67

Jefcoate, C.R. (1978) Measurement of substrate and inhibitor binding to microsomal cytochrome P-450 by optical-difference spectroscopy. *Methods Enzymol.* **52:** 258-279

Johnston, J.B., Ouellet, H., and Ortiz de Montellano, P.R. (2010) Functional redundancy of steroid C26-monooxygenase activity in *Mycobacterium tuberculosis* revealed by biochemical and genetic analyses. *J. Biol. Chem.* **285:** 36352-36360

Joshi, S.M., Pandey, A.K., Capite, N., Fortune, S.M., Rubin, E.J., and Sassetti, C.M. (2006) Characterization of mycobacterial virulence genes through genetic interaction mapping. *Proc. Natl. Acad. Sci. USA* **103:** 11760-11765

Julsing, M.K., Cornelissen, S., Bühler, B., and Schmid, A. (2008) Heme-iron oxygenases: powerful industrial biocatalysts? *Curr. Opin. Chem. Biol.* **12:** 177-186

Jung, C., Ristau, O., and Rein, H. (1991) The high-spin/low-spin equilibrium in cytochrome P-450--a new method for determination of the high-spin content. *Biochim Biophys Acta* **1076:** 130-136

Keizers, P.H., Lussenburg, B.M., de Graaf, C., Mentink, L.M., Vermeulen, N.P., and Commandeur, J.N. (2004) Influence of phenylalanine 120 on cytochrome P450 2D6 catalytic selectivity and regiospecificity: crucial role in 7-methoxy-4-(aminomethyl)-coumarin metabolism. *Biochem. Pharmacol.* **68:** 2263-2271

Kendall, S.L., Burgess, P., Balhana, R., Withers, M., Ten Bokum, A., Lott, J.S., Gao, C., Uhia-Castro, I., and Stoker, N.G. (2010) Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: *kstR* and *kstR2*. *Microbiology* **156:** 1362-1371

Kendall, S.L., Rison, S.C., Movahedzadeh, F., Frita, R., and Stoker, N.G. (2004) What do microarrays really tell us about *M. tuberculosis*? *Trends Microbiol.* **12:** 537-544

Kendall, S.L., Withers, M., Soffair, C.N., Moreland, N.J., Gurcha, S., Sidders, B., Frita, R., Ten Bokum, A., Besra, G.S., Lott, J.S., and Stoker, N.G. (2007) A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol. Microbiol.* **65:** 684-699

Kieslich, K. (1985) Microbial side-chain degradation of sterols. *J. Basic Microbiol.* **25:** 461-474

Kilbane, J.J. 2nd. (2006) Microbial biocatalyst developments to upgrade fossil fuels. *Curr. Opin. Biotechnol.* **17:** 305-314

Kim, D., Kim, Y.-S., Kim, S.-K., Kim, S.W., Zylstra, G.J., Kim, Y.M., and Kim, E. (2002) Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. *Appl. Environ. Microbiol.* **68**: 3270-3278

König, C., Eulberg, D., Gröning, J., Lakner, S., Seibert, V., Kaschabek, S.R., and Schlömann, M. (2004) A linear megaplasmid, p1CP, carrying the genes for chlorocatechol catabolism of *Rhodococcus opacus* 1CP. *Microbiology* **150:** 3075-3087

Knol, J., Bodewits, K., Hessels, G.I., Dijkhuizen, L., and van der Geize, R. (2008) 3-Keto-5alpha-steroid delta1-dehydrogenase from *Rhodococcus erythropolis* SQ1 and its orthologue in *Mycobacterium tuberculosis* H37Rv are highly specific enzymes that function in cholesterol catabolism. *Biochem. J.* **410:** 339-346

Kurakov, V.V., Korovova, Y.N., and Ryzhkova, V.M. (1993) Effect of the initial concentration of beta-sitosterol on its microbiological transformation to androst-4-ene-3,17-dione. *Khim. Farm. Zh.* **27:** 65-67

Lack, N.A., Yam, K.C., Lowe, E.D., Horsman, G.P., Owen, R.L., Sim, E., and Eltis, L.D. (2010) Characterization of a carbon-carbon hydrolase from *Mycobacterium tuberculosis* involved in cholesterol metabolism. *J. Biol. Chem.* **285:** 434-443

Lamb, D.C., Guengerich, F.P., Kelly, S.L., and Waterman, M. R. (2006) Exploiting *Streptomyces coelicolor* A3(2) P450s as a model for application in drug discovery. *Expert Opin. Drug Metab. Toxicol.* **2:** 27-40

Larkin, M.J., Kulakov, L.A., and Allen, C.C. (2005) Biodegradation and *Rhodococcus*--masters of catabolic versatility. *Curr. Opin. Biotechnol.* **16:** 282-290

LeChevalier, H.A. (1989) A practical guide to generic identification of actinomycetes, in *Bergey's Manual of Systematic Bacteriology*, vol. 2 (Sneath, P.H.A., Mair, N.S., Sharpe, M.E., and Holt, J.G., eds) pp. 1458-1506, Williams & Wilkins, Baltimore

Lee, K.U. (1990) Conversion of plant sterol to androst-4-ene-3,17-dione by a mutant of *Mycobacterium* sp. NRRL B-3805. *Kor. J. Microbiol.* **28:** 351-363

Lee, C.Y., Chen, C.D., and Liu, W.H. (1993) Production of androsta-l,4-diene-3,17-dione from cholesterol using two-step microbial transformation. *Appl. Microbiol. Biotechnol.* **38:** 447-452

Lee, C., and Liu, W. (1992) Production of androsta-1,4-diene-3,17-dione from cholesterol using immobilized growing cells of *Mycobacterium* sp. NRRL B-3683 adsorbed on solid carriers. *Appl. Microbiol. Biotechnol.* **36:** 598-603

Levy, H.R., and Talalay, P. (1959a) Bacterial oxidation of steroids. I: ring A dehydrogenations by intact cells. *J. Biol. Chem.* **234:** 2009-2013

Levy, H.R., and Talalay, P. (1959b) Bacterial oxidation of steroids. II: studies on the enzymatic mechanism of ring A dehydrogenation. *J. Biol. Chem.* **234:** 2014-2021

Li, L, Bannantine, J.P., Zhang, Q., Amonsin, A., May, B.J., Alt, D., Banerji, N., Kanjilal, S., and Kapur, V. (2005) The complete genome sequence of *Mycobacterium avium* subspecies paratuberculosis. *Proc. Natl. Acad. Sci. USA* **102:**12344-12349

Liu, W.H., Kuo, C.W., Wu, K.L., Lee, C.Y., and Hsu,W.Y. (1994) Transformation of cholesterol to testosterone by *Mycobacterium* sp. *J. lnd. Microbiol.* **13:** 167-171

Liu, W., and Lee, C.Y. (1992) Production of androst-4-ene-3,17-dione from cholesterol by *Mycobacterium* sp. in a synthetic medium. *Zhongguo Nongye Hyaxue Huizhi* **30:** 52-58

Liu, W.H., and Lo, C.K. (1997) Production of testosterone from cholesterol using a single-step microbial transformation of *Mycobacterium* sp. *J. Ind. Microbiol. Biotechnol.* **19:** 269-272

Liu, P.T., Stenger, S., Li, H., Wenzel, L., Tan, B.H., Krutzik, S.R., Ochoa, M.T., Schauber, J., Wu, K., Meinken, C., Kamen, D.L., Wagner, M., Bals, R., Steinmeyer, A., Zügel, U., Gallo, R.L., Eisenberg, D., Hewison, M., Hollis, B.W., Adams, J.S., Bloom, B.R., and Modlin, R.L. (2006) Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. *Science* **311:** 1770-1773

Lo, CK., Pan, C.P., and Liu, W.H. (2002) Production of testosterone from phytosterol using a single-step microbial transformation by a mutant of *Mycobacterium* sp. *J. Ind. Microbiol. Biotechnol.* **28:** 280-283

Lorenz, P., and Eck, J. (2005) Metagenomics and industrial applications. *Nat. Rev. Microbiol.* **3:** 510-516

Lussenburg, B.M., Babel, L.C., Vermeulen, N.P., and Commandeur, J.N. (2005) Evaluation of alkoxyresorufins as fluorescent substrates for cytochrome P450 BM3 and site-directed mutants. *Anal. Biochem.* **341:** 148-155

Lynd, L.R., Weimer, P.J., van Zyl, W.H., and Pretorius, I.S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66:** 506-577

MacLachlan, J., Wotherspoon, A.T., Ansell, R.O., and Brooks, C.J. (2000) Cholesterol oxidase: sources, physical properties and analytical applications. *J. Steroid Biochem. Mol. Biol.* **72:**169-195

Mahato, S.B., and Garai, S. (1997) Advances in microbial steroid biotransformation. *Steroids* **62:** 332-345

Mallonee, D.H., Adams, J.L., and Hylemon, P.B. (1992) The bile acid-inducible *baiB* gene from *Eubacterium* sp. strain VPI 12708 encodes a bile acid-coenzyme A ligase. *J. Bacteriol.* **174:** 2065-2071

Marchler-Bauer, A., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He, S., Hurwitz, D.I., Jackson, J.D., Jacobs, A.R., Lanczycki, C.J., Liebert, C.A., Liu, C., Madej, T., Marchler, G.H., Mazumder, R., Nikolskaya, A.N., Panchenko, A.R., Rao, B.S., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Vasudevan, S., Wang, Y., Yamashita, R.A., Yin, J.J., and Bryant, S.H. (2003) CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **31:** 383-387

Marsheck, W.J., Kraychy, S., and Muir, R.D. (1972) Microbial degradation of sterols. *Appl. Microbiol.* **23:** 72-77

Martin, C.K.A. (1977) Microbial cleavage of sterol side chains. *Adv. Appl. Microbiol.* **22:** 29-58

Martin, C.K.A., and Wagner, F. (1976) Microbial transformation of beta-sitosterol by *Nocardia* sp. M29. *Eur. J. Appl. Microbiol.* **2:** 243-255

Martínková, L., Uhnáková, B., Pátek, M., Nesvera, J., and Kren V. (2009) Biodegradation potential of the genus *Rhodococcus*. *Environ. Int.* **35:** 162-177

Masai, E., Yamada, A., Healy, J.M., Hatta, T., Kimbara, K., Fukuda, M., and Yano, K. (1995) Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol.* **61:** 2079-2085

Mast, N., Murtazina, D., Liu, H., Graham, S.E., Bjorkhem, I., Halpert, J.R., Peterson, J., and Pikuleva, I.A. (2006) Distinct binding of cholesterol and 5beta-cholestane-3alpha,7alpha,12alpha-triol to cytochrome P450 27A1: evidence from modeling and site-directed mutagenesis studies. *Biochemistry* **45:** 4396-4404

Masuda, S., Prosser, D.E., Guo, Y.D., Kaufmann, M., and Jones, G. (2007) Generation of a homology model for the human cytochrome P450, CYP24A1, and the testing of putative substrate binding residues by site-directed mutagenesis and enzyme activity studies. *Arch. Biochem. Biophys.* **460:** 177-191

Mathieu, J., Mohn, W.W., Eltis, L.D., Leblanc, J., Stewart, G., Dresen, C., Okamoto, K., and Alvarez, P.J. (2010) 7-Ketocholesterol catabolism by *Rhodococcus jostii* RHA1. *Appl. Environ. Microbiol*. **76:** 352-355

Mathur, S., Bhatia, M.C., and Mathur, S.N. (1992) Biotransformation of beta-sitosterol into 17-ketosteroids by some strains of *Arthrobacter.* in *Role of biotechnology in agriculture* (Prasad, B.N., Ghimire, G.P.S., and Agrawal, V.P., eds) pp. 99-102, Oxford & IBH Pub., New Delhi

McLean, K.J., Clift, D., Lewis, D.G., Sabri, M., Balding, P.R., Sutcliffe, M.J., Leys, D., and Munro, A.W. (2006) The preponderance of P450s in the *Mycobacterium tuberculosis* genome. *Trends Microbiol.* **14:** 220-228

McLean, K.J., Lafite, P., Levy, C., Cheesman, M.R., Mast, N., Pikuleva, I.A., Leys, D., and Munro, A.W. (2009) The structure of *Mycobacterium tuberculosis* CYP125: molecular basis for cholesterol binding in a P450 needed for host infection. *J. Biol. Chem.* **284:** 35524-35533

McLeod, M.P., Warren, R.L., Hsiao, W.W., Araki, N., Myhre, M., Fernandes, C., Miyazawa, D., Wong, W., Lillquist, A.L., Wang, D., Dosanjh, M., Hara, H., Petrescu, A., Morin, R.D., Yang, G., Stott, J.M., Schein, J.E., Shin, H., Smailus, D., Siddiqui, A.S., Marra, M.A., Jones, S.J., Holt, R., Brinkman, F.S., Miyauchi, K., Fukuda, M., Davies, J.E., Mohn, W.W., and Eltis, L.D. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc. Natl. Acad. Sci. USA* **103:** 15582-15587

Medema, M.H., Trefzer, A., Kovalchuk, A., van den Berg, M., Müller, U., Heijne, W., Wu, L., Alam, M.T., Ronning, C.M., Nierman, W.C., Bovenberg, R.A., Breitling, R., and Takano, E. (2010) The sequence of a 1.8-mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biol. Evol.* **12:** 212-224

Miclo, A., and Germain, P. (1990) Catabolism of methylperhydroindanedione propionate by *Rhodococcus equi*: evidence of a MEPHIP-reductase activity. *Appl. Microbiol. Biotechnol.* **32:** 594-599

Miclo, A., and Germain, P. (1992) Hexahydroindanone derivatives of steroids formed by *Rhodococcus equi*. *Appl. Microbiol. Biotechnol.* **36:** 456-460

Miyamoto, M., Matsumoto, J., Iwaya, T., and Itagaki, E. (1995) Bacterial steroid monooxygenase catalyzing the Baeyer-Villiger oxidation of C21-ketosteroids from *Rhodococcus rhodochrous*: the isolation and characterization. *Biochim. Biophys. Acta* **1251:** 115-124

Metcalfe, A.C., Krsek, M., Gooday, G.W., Prosser, J.I., and Wellington, E.M. (2002)
Molecular analysis of a bacterial chitinolytic community in an upland pasture. *Appl. Environ. Microbiol.* **68:** 5042-5050

Mindnich, R., Möller, G., and Adamski, J. (2004) The role of 17 beta-hydroxysteroid dehydrogenases. *Mol. Cell. Endocrinol.* **218:** 7-20

Mohn, W.W., van der Geize, R., Stewart, G.R., Okamoto, S., Liu, J., Dijkhuizen, L., and Eltis, L.D. (2008) The actinobacterial *mce4* locus encodes a steroid transporter. *J. Biol. Chem.* **283:** 35368-35374

Molnár, I., Hayashi, N., Choi, K.-P., Yamamoto, H., Yamashita, M., and Murooka, Y. (1993) Bacterial cholesterol oxidases are able to act as flavoprotein-linked ketosteroid monooxygenases that catalyse the hydroxylation of cholesterol to 4-cholest-6-ol-3-one. *Mol. Microbiol.* **7:** 419-428

Morii, S., Sawamoto, S., Yamauchi, Y., Miyamoto, M., Iwami, M., and Itagaki, E. (1999) Steroid monooxygenase of *Rhodococcus rhodochrous*: sequencing of the genomic DNA, and hyperexpression, purification, and characterization of the recombinant enzyme. *J. Biochem.* **126:** 624-631

Murohashi, T., Kondo, E., and Yoshida, K. (1969) The role of lipids in acid-fastness of mycobacteria. *Am. Rev. Respir. Dis.* **99:** 794-798

Murohisa, T., and Iida, M. (1993a) Studies on microbial transformation (XXVI). Microbial degradation of 19-hydroxysterol side-chains. *J. Ferment. Bioeng.* **75:** 13-17

Murohisa, T., and Iida, M. (1993b) Studies on microbial transformation (XXVII). Some new intermediates in microbial side-chain degradation. *J. Ferment. Bioeng.* **76:** 174-177

Murooka, Y., Ishizaki, T., Nimi, O., and Maekawa, N. (1986) Cloning and expression of a *Streptomyces* cholesterol oxidase gene in *Streptomyces lividans* with plasmid pIJ702. *Appl. Environ. Microbiol.* **52:** 1382-1385

Nailor, M.D., and Sobel, J.D. (2009) Antibiotics for gram-positive bacterial infections: vancomycin, teicoplanin, quinupristin/dalfopristin, oxazolidinones, daptomycin, dalbavancin, and telavancin. *Infect. Dis. Clin. North Am.* **23:** 965-982

Nakashima, N., and Tamura, T. (2004) Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. *Appl. Environ. Microbiol.* **70:** 5557-5568

Navas, J., González-Zorn, B., Ladrón, N., Garrido, P., and Vázquez-Boland, J.A. (2001) Identification and mutagenesis by allelic exchange of *choE*, encoding a cholesterol oxidase from intracellular pathogen *Rhodococcus equi*. *J. Bacteriol.* **183:** 4796-4805

Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C., and Nebert, D.W. (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6:** 1-42

Nesbitt, N.M., Yang, X., Fontán, P., Kolesnikova, I., Smith, I., Sampson, N.S., and Dubnau, E. (2010) A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infect. Immun.* **78:** 275-282

Neu, T.R. (1996) Significance of bacterial surface-active compounds in interaction of bacteria with interfaces. *Microbiol. Rev.* **60:** 151-166

Nilekani, S., and SivaRaman, C. (1983) Purification and properties of citrate lyase from *Escherichia coli*. *Biochemistry* **22:** 4657-4663

Oda, S. (1994) Manufacture of 1,4-androstadiene-3,17-dione with immobilized microorganism. Japanese Patent 0600, 095 (Kansai Paint Co. Ltd)

Omura, T., and Sato, R. (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J. Biol. Chem.* **239:** 2370-2378

Ouellet, H., Guan, S., Johnston, J.B., Chow, E.D., Kells, P.M., Burlingame, A.L., Cox, J.S., Podust, L.M., and de Montellano, P.R. (2010) *Mycobacterium tuberculosis* CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one. *Mol. Microbiol.* **77:** 730-742

Ouellet, H., Lang, J., Couture, M., and Ortiz de Montellano, P.R. (2009) Reaction of *Mycobacterium tuberculosis* cytochrome P450 enzymes with nitric oxide. *Biochemistry* **48:** 863-872

Ourisson, G., Albrecht, P., and Rohmer, M. (1979) The hopanoids: palaeochemistry and biochemistry of a group of natural products. *Pure Appl. Chem.* **51:** 709-729

Paje, M.L., Neilan, B.A., and Couperwhite, I. (1997) A *Rhodococcus* species that thrives on medium saturated with liquid benzene. *Microbiology* **143:** 2975-2981

Pandey, A.K., and Sassetti, C.M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci. USA* **105:** 4376-4380

Parikh, A., Gillam, E.M., and Guengerich, F.P. (1997) Drug metabolism by *Escherichia coli* expressing human cytochromes P450. *Nat. Biotechnol.* **15:** 784-788

Patrauchan, M.A., Florizone, C., Dosanjh, M., Mohn, W.W., Davies, J., and Eltis, L.D. (2005) Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. *J. Bacteriol.* **187:** 4050-4063

Pei, Z., Oey, N.A., Zuidervaart, M.M., Jia, Z., Li, Y., Steinberg, S.J., Smith, K.D., and Watkins, P.A. (2003) The acyl-CoA synthetase "bubblegum" (lipidosin): further characterization and role in neuronal fatty acid beta-oxidation. *J. Biol. Chem.* **278:** 47070-47078

Peretó, J., López-García, P., and Moreira, D. (2005) Phylogenetic analysis of eukaryotic thiolases suggests multiple proteobacterial origins. *J. Mol. Evol.* **61:** 65-74

Pethe, K., Swenson, D.L., Alonso, S., Anderson, J., Wang, C., and Russell, D.G. (2004) Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc. Natl. Acad. Sci. USA* **101:** 13642-13647

Petrusma, M., Dijkhuizen, L., and van der Geize, R. (2009) *Rhodococcus rhodochrous* DSM43269 3-ketosteroid 9alpha-hydroxylase, a two-component iron-sulfur-containing monooxygenase with subtle steroid substrate specificity. *Appl. Environ. Microbiol.* **75:** 5300-5307

Peyron, P., Bordier, C., N'Diaye, E.N., and Maridonneau-Parini, I. (2000) Nonopsonic phagocytosis of *Mycobacterium kansasii* by human neutrophils depends on cholesterol and is mediated by CR3 associated with glycosylphosphatidylinositol-anchored proteins. *J. Immunol.* **165:** 5186-5191

Pikuleva, I.A. (2006) Cholesterol-metabolizing cytochromes P450. *Drug Metab. Dispos.* **34:** 513-520

Pizzul, L., Sjögren, A., Castillo, M. del P., and Stenström, J. (2007) Degradation of polycyclic aromatic hydrocarbons in soil by a two-step sequential treatment. *Biodegradation* **18:** 607-166

Plesiat, P., Grandguillot, M., Harayama, S., Vragar, S., and Michel-Briand, Y. (1991) Cloning, sequencing, and expression of the *Pseudomonas testosteroni* gene encoding 3-oxosteroid delta1-dehydrogenase. *J. Bacteriol.* **173:** 7219-7227

Prome, D., Lacave, C., Monsarrat, B., David, H., and Prome, J.C. (1983) Conversion of sterols and triterpenes by mycobacteria. I Formation of progesterone and 1-dehydroprogesterone from *Mycobacterium aurum*, strain A^+. *Biochim. Biophys. Acta* **753:** 60-64

Reiche, R., Heller, I., Hoerhold, C., and Gotlsehaldf, B. (1992) Manufacture of androsta-l,4-diene-3,17-dione from sterols with *Mycobacterium.* (East) German Patent 300, 364 (VEBJENA PHARM)

Rengarajan, J., Bloom, B.R., and Rubin, E.J. (2005) Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proc. Natl. Acad. Sci. USA* **102:** 8327-8332

Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C., Withers, S.T., Shiba, Y., Sarpong, R., and Keasling, J.D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440:** 940-943

Ropp, J.D., Gunsalus, I.C., and Sligar, S.G. (1993) Cloning and expression of a member of a new cytochrome P-450 family: cytochrome P-450lin (CYP111) from *Pseudomonas incognita*. *J. Bacteriol.* **175:** 6028-6037

Rosłoniec, K.Z., Wilbrink, M.H., Capyk, J.K., Mohn, W.W., Ostendorf, M., van der Geize, R., Dijkhuizen, L., and Eltis, L.D. (2009) Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in *Rhodococcus jostii* RHA1. *Mol. Microbiol.* **74:** 1031-43

Roy, P.K., Khan, A.W., Kumar, J., Chopra, S.D.K., and Basu, S.K. (1992) Steroid transformation in a laboratory-scale glass airlift fermentor. *World J. Microbiol. Biotechnol.* **8:** 399-401

Sakai, M., Masai, E., Asami, H., Sugiyama, K., Kimbara, K. and Fukuda, M. (2002) Diversity of 2,3-dihydroxybiphenyl dioxygenase genes in a strong PCB degrader, *Rhodococcus* sp. strain RHA1. *J. Biosci. Bioeng.* **93:** 421-427

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48:** 77-84

Sassetti, C.M., and Rubin, E.J. (2003) Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. USA* **100:** 12989-12994

Schmid, A., Dordick, J.S., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B. (2001) Industrial biocatalysis today and tomorrow. *Nature* **409:** 258-268

Schmidt, D.D., Frommer, W., Junge, B., Muller, L., Wingender, W., Truscheit, E., and Schafer, D. (1977) Alpha-glucosidase inhibitors: new complex oligosaccharides of microbial origin. *Naturwissenschaften* **64:** 535-536

Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C., and Schoolnik, G.K. (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J. Exp. Med.* **198:** 693-704

Seedorf, U., Brysch, P., Engel, T., Schrage, K., and Assmann, G. (1994) Sterol carrier protein X is peroxisomal 3-oxoacyl coenzyme A thiolase with intrinsic sterol carrier and lipid transfer activity. *J. Biol. Chem.* **269:** 21277-21283

Seidel, L., and Hoerhold, C. (1992) Selection and characterization of new microorganisms for the manufacture of 9-OH-AD from sterols. *J. Basic Microbiol.* **32:** 49-55

Seto, M., Kimbara, K., Shimura, M., Hatta, T., Fukuda, M., and Yano, K. (1995) A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol.* **61:** 3353-3358

Shah, N.S., Wright, A., Bai, G.H., Barrera, L., Boulahbal, F., Martín-Casabona, N., Drobniewski, F., Gilpin, C., Havelková, M., Lepe, R., Lumb, R., Metchock, B., Portaels, F., Rodrigues, M.F., Rüsch-Gerdes, S., van Deun, A., Vincent, V., Laserson, K., Wells, C., and Cegielski, J.P. (2007) Worldwide emergence of extensively drug-resistant tuberculosis. *Emerg. Infect. Dis.* **13:** 380-387

Shi, J., Chu, Z., Ju, Y., Mo, G., and Cheng, W. (1992) Degradation of cholesterol to androsta-1,4-diene-3,17-dione by *Rhodococcus corallina. Zhongguo Yiyao Gongye Zazhi* **23:** 204-207

Shukla, A., Patil, S., and Bharti, S. (1992) Microbial conversion of solasodine to 4-androstene-3,17-dione (AD), a key intermediate for androgen synthesis. *Lett. Appl. Microbiol.* **15:** 86-88

Sih, C.J. (1962) Mechanisms of steroid oxidation by microorganisms. *Biochim. Biophys. Acta.* **62:** 541-547

Sih, C.J., and Bennet, R.E. **(**1962) Steroid 1-dehydrogenase of *Nocardia restrictus*. *Biochim. Biophys. Acta* **56**: 587-592

Sih, C.J., Wang, K.C., and Tai, H.H. (1968a) Mechanism of steroid oxidation by microorganisms XIII. C22 acid intermediates in the degradation of the cholesterol side chain. *Biochemistry* **7:** 796-807

Sih, C.J., Tai, H.H., and Tsong, Y.Y. (1967) The mechanism of microbial conversion of cholesterol into 17-keto steroids. *J. Am. Chem. Soc.* **89:** 1957-1958

Sih, C.J., Tai, H.H., Tsong, Y.Y., Lee, S.S., and Coombe, R.G. (1968b) Mechanisms of steroid oxidation by microorganisms. XIV. Pathway of cholesterol side-chain degradation. *Biochemistry* **7:** 808-818

Sim, L., Willemsma, C., Mohan, S., Naim, H.Y., Pinto, B.M., and Rose, D.R. (2010) Structural basis for substrate selectivity in human maltase-glucoamylase and sucrase-isomaltase N-terminal domains. *J. Biol. Chem.* **285:** 17763-17770

Smith, M., Zahnley, J., Pfeifer, D., and Goff, D. (1993) Growth and cholesterol oxidation by *Mycobacterium* species in Tween 80 medium. *Appl. Environ. Microbiol.* **59:** 1425-1429

Söhngen, N.L. (1913) Benzin, Petroleum, Paraffinöil und Paraffin als Kohlenstoff- und Energiequelle für Mikroben. *Zbl. Bakteriol. Parasitenk. Abt. II* **37:** 595-609

Sojo, M., Bru, R., Lopez-Molina, D., Garcia-Carmona, F., and Arguelles, J.C. (1997) Cell-linked and extracellular cholesterol oxidase activities from *Rhodococcus erythropolis*. Isolation and physiological characterization. *Appl. Microbiol. Biotechnol.* **47:** 583-589

Song, J., Wadhwa, L., Bejjani, B.A., and O'Brien, W.E. (2003) Determination of 3-keto-4-ene steroids and their hydroxylated metabolites catalyzed by recombinant human cytochrome P450 1B1 enzyme using gas chromatography-mass spectrometry with trimethylsilyl derivatization. *J. Chromatogr. B. Analyt. Technol. Biomed. Life. Sci.* **791:** 127-135

Spassov, G., Pramatarova, V., Vlahov, R., and Reinhold, G. (1993) Bioconversion of sitosterol to androstene-dione by mutants of *Mycobacterium vaccae. Dokl. Bulg. Acad. Nauk.* **46:** 123-126

Srinivasan, A., Dick, J.D., and Perl, T.M. (2002) Vancomycin resistance in staphylococci. *Clin. Microbiol. Rev.* **15:** 430-438

Srivastava, A., and Patil, S. (1994) Investigation of some physicochemical parameters involved in the biotransformation of cholesterol to 17-ketosteroids. *J. Microb. Biotechnol.* **9:** 101-112

Stadtman, T.C., Cherkes, A., and Anfinsen, C.N. (1954) Studies on the microbial degradation of cholesterol. *J. Biol. Chem.* **206:** 511-523

Stecker, C., Johann, A., Herzberg, C., Averhoff, B., and Gottschalk, G. (2003) Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. *J. Bacteriol.* **185:** 5269-5274

Stegink, L.D., and Coon, M.J. (1968) Stereospecificity and other properties of highly purified beta-hydroxy-beta-methylglutaryl coenzyme A cleavage enzyme from bovine liver. *J. Biol. Chem.* **243:** 5272-5279

Storbeck, K.H., Swart, P., and Swart, A.C. (2007) Cytochrome P450 side-chain cleavage: insights gained from homology modeling. *Mol. Cell Endocrinol.* **265-266:** 65-70

Szentirmai, A. (1990) Microbial physiology of sidechain degradation of sterols. *J. Ind. Microbiol.* **6:** 101-116

Strijewski, A. (1982) The steroid-9alpha-hydroxylation system from *Nocardia* species. *Eur. J. Biochem.* **128:** 125-135

Sutcliffe, I.C. (1998) Cell envelope composition and organisation in the genus *Rhodococcus*. *Antonie Van Leeuwenhoek* **74:** 49-58

Szentirmai, A. (1990) Microbial physiology of side chain degradation of sterols. *J. Ind. Microbiol. Biotechnol.* **6:** 101-116

Takai, S., Sekizaki, T., Ozawa, T., Sugawara, T., Watanabe, Y., and Tsubaki, S. (1991) Association between a large plasmid and 15- to 17-kilodalton antigens in virulent *Rhodococcus equi*. *Infect. Immun.* **59:** 4056-4060

Takai, S., Watanabe, Y., Ikeda, T., Ozawa, T., Matsukura, S., Tamada, Y., Tsubaki, S., and Sekizaki, T. (1993) Virulence-associated plasmids in *Rhodococcus equi*. *J. Clin. Microbiol.* **31:** 1726-1729

Takeuchi, H., Chen, J.H., Jenkins, J.R., Bun-Ya, M., Turner, P.C., and Rees, H.H. (2004) Characterization of a sterol carrier protein 2/3-oxoacyl-CoA thiolase from the cotton leafworm (*Spodoptera littoralis*): a lepidopteran mechanism closer to that in mammals than that in dipterans. *Biochem. J.* **382:** 93-100

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24:** 1596-1599

Tenneson, M.E., Baty, J.D., Bilton, R.F., and Mason, A.N. (1979) The degradation of cholic acid by *Pseudomonas* sp. N.C.I.B. 10590. *Biochem. J.* **184:** 613-618

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673-4680

Ullah, A.J., Murray, R.I., Bhattacharyya, P.K., Wagner, G.C., and Gunsalus, I.C. (1990) Protein components of a cytochrome P-450 linalool 8-methyl hydroxylase. *J. Biol. Chem.* **265:** 1345-1351

Trivedi, O.A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., and Gokhale, R.S. (2004) Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. *Nature* **428:** 441-445

Vaillancourt, F.H., Haro, M.A., Drouin, N.M., Karim, Z., Maaroufi, H., and Eltis, L.D. (2003) Characterization of extradiol dioxygenases from a polychlorinated biphenyl-degrading strain that possess higher specificities for chlorinated metabolites. *J. Bacteriol.* **185:** 1253-1260

van der Geize, R., de Jong, W., Hessels, G.I., Grommen, A.W., Jacobs, A.A., and Dijkhuizen, L. (2008a) A novel method to generate unmarked gene deletions in the intracellular pathogen *Rhodococcus equi* using 5-fluorocytosine conditional lethality. *Nucleic Acids Res.* **36:** e151

van der Geize, R., and Dijkhuizen, L. (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Curr. Opin. Microbiol.* **7:** 255-261

van der Geize, R., Hessels, G.I., and Dijkhuizen, L. (2002a) Molecular and functional characterization of the kstD2 gene of *Rhodococcus erythropolis* SQ1 encoding a second 3-ketosteroid delta1-dehydrogenase isoenzyme. *Microbiology* **148:** 3285-3292

van der Geize, R., Hessels, G.I., and Dijkhuizen, L. (2009) Method for the production of modified steroid degrading microorganisms and their use. *USPTO Application 20090186390*

van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., and Dijkhuizen, L. (2001) Unmarked gene deletion mutagenesis of *kstD*, encoding 3-ketosteroid delta1-dehydrogenase, in *Rhodococcus erythropolis* SQ1 using *sac*B as counter-selectable marker. *FEMS Microbiol. Lett.* **205:**197-202

van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., and Dijkhuizen, L. (2002b) Molecular and functional characterization of *kshA* and *kshB*, encoding two components of 3-ketosteroid 9alpha-hydroxylase, a class IA monooxygenase, in *Rhodococcus erythropolis* strain SQ1. *Mol. Microbiol.* **45:** 1007-1018

van der Geize, R., Hessels, G.I., van Gerwen, R., Vrijbloed, J.W., van der Meijden, P., and Dijkhuizen, L. (2000) Targeted disruption of the *kstD* gene encoding a 3-ketosteroid delta1-dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. *Appl. Envir. Microbiol.* **66:** 2029-2036

van der Geize, R., Hessels, G.I., Nienhuis-Kuiper, M., and Dijkhuizen, L. (2008b) Characterization of a second *Rhodococcus erythropolis* SQ1 3-ketosteroid 9alpha-hydroxylase activity comprising a

terminal oxygenase homologue, KshA2, active with oxygenase-reductase component KshB. *Appl. Environ. Microbiol.* **74:** 7197-7203

van der Geize, R., van der Meijden, P., Hessels, G.I., and Dijkhuizen, L. (2003) Identification of 3-ketosteroid 9alpha-hydroxylase genes and microorganisms blocked in 3-ketosteroid 9alpha-hydroxylase activity. *Patent WO 03/070925. Akzo Nobel N.V. (The Netherlands)*

van der Geize, R., Yam, K., Heuser, T., Wilbrink, M.H., Hara, H., Anderton, M.C., Sim, E., Dijkhuizen, L., Davies, J.E., Mohn, W.W., and Eltis, L.D. (2007) A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc. Natl. Acad. Sci. USA* **104:** 1947-1952

van Wezel, G.P., McKenzie, N.L., and Nodwell, J.R. (2009) Chapter 5. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics. *Methods Enzymol.* **458:** 117-141

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., van Sinderen, D. (2007) Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev.* **71:** 495-548

Vert, G., Nemhauser, J.L., Geldner, N., Hong, F., and Chory, J. (2005). Molecular mechanisms of steroid hormone signaling in plants. *Annu. Rev. Cell. Dev. Biol.* **21:** 177-201

Volkman, J. K. (2003) Sterols in microorganisms. *Appl. Microbiol. Biotechnol.* **60:** 495-506

Vorbrodt, H.M., Adam, G., Porzel, A., Hoerhold, C., Daenhardt, S., and Boehme, K.H. (1991) Microbial degradation of 2alpha,3alpha-dihydroxy-5alpha-cholestan-6-one by *Mycobacterium vaccae. Steroids* **56:** 586-588

Vrielink, A., and Ghisla, S. (2009) Cholesterol oxidase: biochemistry and structural features. *FEBS J.* **276:** 6826-6843

Wang, K.C., Chiang, N., and Su, C.C. (1994) Microbial transformation of 5beta-steroids with *Mycobacterium* sp. (NRRL B-3805). *Chin. Pharm. J. (Taipei)* **46:** 227-232

Wang, K.C., You, B.-J., Yan, J.-L., and Lee, S.-S. (1995) Microbial transformation of lanosterol derivatives with *Mycobacterium* sp. NRRL B-3805. *J. Nat. Prod.* **58:** 1222-1227

Watanabe, K., Aihara, H., Tachi, N., and Nakamura, R. (1987) Degradations of 4-cholesten-3-one and 1,4-androstadiene-3,17-dione by cholesterol-degrading bacteria. *J. Appl. Bacteriol.* **62:** 151-155

Weber, A., Kennecke, M., and Neef, G. (1992) 20-Methyl-5,7-pregnadiene-3beta,2l-diol derivatives and their manufacture with *Mycobacterium.* World Patent 0903, 465 (Schering A-G)

Wilbrink, M.H., Dijkhuizen, L., and van der Geize, R. FadD19 of *Rhodococcus rhodochrous* DSM43269: a steroid-CoA ligase essential for degradation of C24-branched sterol side chains. Submitted

Williams, J.R., Chai, D., and Wright, D. (2002) Synthesis of (25*R*)-26-hydroxycholesterol. *Steroids* **67:** 1041-1044

Wilmańska, D., Dziadek, J., Sajduda, A., Milczarek, K., Jaworski, A., and Murooka, Y. (1995) Identification of cholesterol oxidase from fast-growing mycobacterial strains and *Rhodococcus* sp. *J. Ferment. Bioeng.* **79:** 119-124

Wolfe, S.L. (1993) Molecular and Cellular Biology. Wadsworth Publishing Company

Wood, M.J. (1996) The comparative efficacy and safety of teicoplanin and vancomycin. *J. Antimicrob. Chemother.* **37:** 209-222

World Health Organization (2005) Global TB Fact Sheet 104 (WHO, Geneva)

Wovcha, M.G., Antosz, F.J., Knight, J.C., Kominek, L.A., and Pyke, T.R. (1978) Bioconversion of sitosterol to useful steroidal intermediates by mutants of *Mycobacterium fortuitum*. *Biochim. Biophys. Acta* **531:** 308-321

Yam, K.C., D'Angelo, I., Kalscheuer, R., Zhu, H., Wang, J.X., Snieckus, V., Ly, L.H., Converse, P.J., Jacobs, W.R. Jr., Strynadka, N., and Eltis, L.D. (2009) Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of *Mycobacterium tuberculosis*. *PLoS Pathog.* e1000344

Yamada, H., and Nagasawa, T. (1994) Process for biological production of amides with *R. rhodochrous* J1. US Patent 5334519

Yang, X., Dubnau, E., Smith, I., and Sampson, N.S. (2007) Rv1106c from *Mycobacterium tuberculosis* is a 3β-hydroxysteroid dehydrogenase. *Biochemistry* **46:** 9058-9067

Yang, X., Nesbitt, N.M., Dubnau, E., Smith, I., and Sampson, N.S. (2009) Cholesterol metabolism increases the metabolic pool of propionate in *Mycobacterium tuberculosis*. *Biochemistry* **48:** 3819-3821

Zaretskaya, I.I, Kogan, L.M., Tikhomirova, O.B., Sis, Jh.D., Wulfson, N. S., Zaretskii, V. I., Zaikin, V.G., Skryabin, G.K., and Torgov, I.V. (1968) Microbial hydroxylation of the cholesterol side chain. *Tetrahedron* **24:** 1595-1600

Zhang, Y. (2005) The magic bullets and tuberculosis drug targets. *Annu. Rev. Pharmacol. Toxicol.* **45:** 529-564

Zhang, L.Q., Bian, E.P., and Wang, Y. (1992) Side-chain cleavage of sterols by *Mycobacterium* sp. MI2. *Yao Xue Xue Bao* **27:** 903-907

DANKWOORD

Eindelijk is het dan zover, ik ben aangekomen bij het laatste aspect van het schrijven van dit proefschrift. De tijd lijkt omgevlogen en ik kan terugkijken op een mooie en leerzame tijd als AIO waarin ook privé het nodige is gebeurd. Op deze pagina's wil ik graag de mensen bedanken die op welke manier dan ook hebben bijgedragen aan de totstandkoming van dit proefschrift.

Allereerst mijn promotor, Lubbert. Bedankt dat ik in jouw onderzoeksgroep ontzettend leuk onderzoek heb mogen doen. Je hebt altijd een drukke agenda gehad maar voor vragen kon ik altijd bij je terecht en replies op e-mails kwamen altijd spoedig (meestal rond middernacht) binnen. Hoewel het project niet altijd even soepel verliep bleef je positief over de behaalde vorderingen en hield je vooral de grote lijnen en de mogelijkheden in het vizier waar ik me nog wel eens verloor in de details. Daar heb ik veel van geleerd. Ik wil je dan ook heel hartelijk bedanken voor het in mij gestelde vertrouwen.
Daarnaast wil ik mijn copromotor en begeleider Robert bedanken. Je enthousiasme voor wetenschap in het algemeen en *Rhodococcus* en steroïde afbraak in het bijzonder was aanstekelijk en erg motiverend. Ik heb heel veel van je geleerd en ben onder de indruk van wat je met je onderzoekslijn allemaal hebt bereikt. Ook je kritische houding met betrekking tot het schrijven van wetenschappelijk artikelen was erg leerzaam en behulpzaam, de manuscripten zijn er stukken leesbaarder van geworden. Ik vond het erg prettig om met je samen te werken en een kantoor met je te delen en er zitten voor de toekomst nog een aantal leuke dingen in "de pijplijn".

I would like to thank the other members of the "*Rhodococcus* team"; Kamila, Jan, Mirjan, Gerda and Manny for all their help, discussions, concerns, laughs and the great time spent in D202 and the big lab. Furthermore, I would like to thank all the other members of MicPhys for the great atmosphere in (and outside) the lab, the fruitful discussions, help with various problems and of course for sharing ups and downs that go together with doing research and life in general. Quite a number of people have come and gone during my stay in the group and I will make an attempt to mention all: Alle, Ana, Andriy, Bave, Coenie, Davide, Dennis, Elena, Eriko, Evelien, Fean, Gert-Jan, Hans K., Hans L., Harm, Ilse, Irma, Jeroen, Jolanda, Justyna, Karina, Lara, Lukasz, Marc, Marco, Mark, Marnix, Marta, Munir, Nai-Hua, Nynke, Peter S., Peter de V., Pieter G., Pieter S., Rachel, Ronan, Sander, Slavko, Stef, Theo, Thijs, Ulrike, Vincent, Wieger, Wouter and of course all the students who enlivened our group.
De dames van het secretariaat (Bea, Manon, Marga, Jasmijn) wil ik bedanken voor de ondersteuning met alle administratieve zaken. Anna en Mozes zorgden dat er altijd voldoende schoon glaswerk was en dat pipetten gekalibreerd werden, waarvoor mijn dank.

I am greatly indebted to the reading committee for their time to evaluate my thesis.
Furthermore I would like to thank prof. Bill Mohn for his hospitality during my stay in his lab at the University of British Columbia, Vancouver. I definitely had a very pleasant stay in Vancouver and learned a lot. The decision to perform DNA-microarrays experiments with steroid grown RHA1 cells ultimately proved golden. The resulting collaboration on

subsequent papers was very pleasant and stimulating for me and I would especially like to acknowledge prof. Lindsay Eltis for his great contributions.

De contactpersonen van onze industriële partners, Peter van der Meijden en Hans Stoelwinder (Organon) en Paul Link en Berthus Bout (Solvay) wil ik hartelijk bedanken voor hun bijdragen aan de discussies tijdens projectmeetings en gastvrijheid in respectievelijk Oss en Weesp. Ook wil ik prof. Dick Janssen en Roga Kembaren bedanken voor hun inbreng in het IBOS project.

Het (af)schrijven van een proefschrift naast een nieuwe functie is zeker niet altijd makkelijk. Daarom wil ik graag prof. Hans Kamerling bedanken voor zijn begrip en vertrouwen gedurende mijn vervolgaanstelling.

Ik wil graag mijn paranimfen Jeroen en Mirjan bedanken dat ze mij terzijde willen staan op de grote dag. Jeroen, jij hebt al ervaring met promoveren en Mirjan, jou wens ik veel succes met de laatste loodjes van het schrijven van jouw proefschrift.

Vrienden en familie waren erg belangrijk voor een luisterend oor, om dingen te relativeren en natuurlijk voor de broodnodige ontspanning. Daarom wil ik mijn vrienden Ilja, Bas, Lisette, Sander, Marieke, Selmar en Joanna bedanken voor alle keren dat we hebben afgesproken in o.a. Groningen, Kampen, Rotterdam, Nijmegen en Leuven. Zo ziet een Groninger nog wat van Nederland (en België). Bas en Sander, jullie zijn mij al succesvol voorgegaan met promoveren en nu zijn we binnen vijf maanden alle drie gepromoveerd, wie had dat gedacht? Ilja wil ik speciaal bedanken voor zijn inzet met allerlei klussen (verhuizen, computers enz.), zijn loyaliteit en gezelligheid.

Natuurlijk wil ik graag mijn ouders bedanken voor hun onvoorwaardelijke steun en geloof in mij. Hans, ik wil je bedanken voor je geweldige vriendschap. Het is geweldig om een broer te hebben om zoveel mee te kunnen delen, van films kijken, klussen, lachen, gamen, koffie drinken tot natuurlijk muziek maken/luisteren en voor het ontdekken van nieuwe bands. Hopelijk heb ik vanaf nu meer tijd om met jou (en Josine) door te brengen!
Ook mijn schoonouders en de rest van mijn (schoon)familie wil ik bedanken voor hun interesse en belangstelling en hulp met allerhande zaken.

Als laatste wil ik mijn gezin bedanken voor alle liefde en steun gedurende alle jaren. Jessica, je hebt met de komst van Anne en Eva je eigen (studie)plannen drastisch omgegooid om er helemaal te zijn voor de kinderen. Ik heb grote bewondering voor de manier waarop je dat al die tijd hebt gedaan. Bedankt ook dat je altijd positief bent gebleven, ook als ik het soms even niet meer zag zitten. Verder ben ik heel blij dat je me hebt geholpen met de lay-out en het ontwerpen van de omslag, jouw creativiteit kent geen grenzen.
Ik prijs me gelukkig met zo'n lieve vrouw en twee fantastische dochters en ik houd van jullie!

LIST OF PUBLICATIONS

Leemhuis, H., Rozeboom, H.J., Wilbrink, M., Euverink, G.J., Dijkstra, B.W., and Dijkhuizen, L. (2003) Conversion of cyclodextrin glycosyltransferase into a starch hydrolase by directed evolution: the role of alanine 230 in acceptor subsite +1. *Biochemistry* **42:** 7518-7526

van der Geize, R., Yam, K., Heuser, T., Wilbrink, M.H., Hara, H., Anderton, M.C., Sim, E., Dijkhuizen, L., Davies, J.E., Mohn, W.W., and Eltis, L.D. (2007) A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc. Natl. Acad. Sci. USA* **104:** 1947-1952

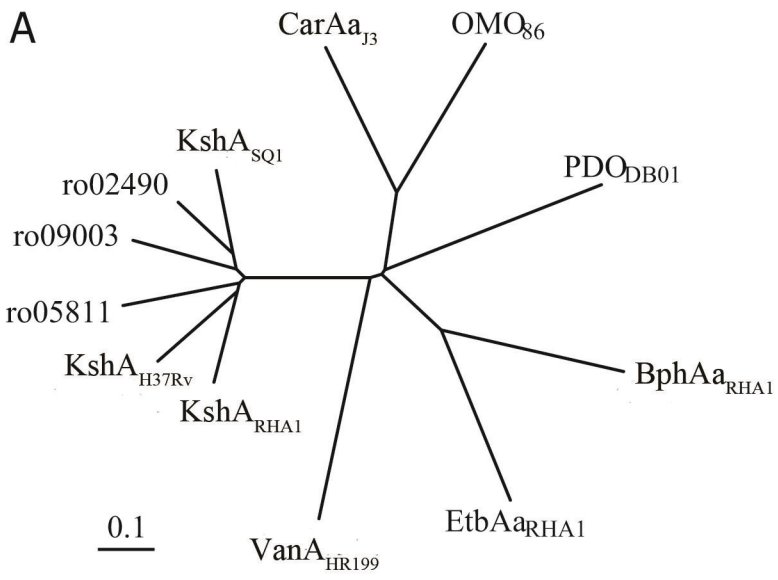
Rosłoniec, K.Z.[#], Wilbrink, M.H.[#], Capyk, J.K., Mohn, W.W., Ostendorf, M., van der Geize, R., Dijkhuizen, L., and Eltis, L.D. (2009) Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in *Rhodococcus jostii* RHA1. *Mol. Microbiol.* **74:** 1031-1043
# These authors contributed equally to this work.

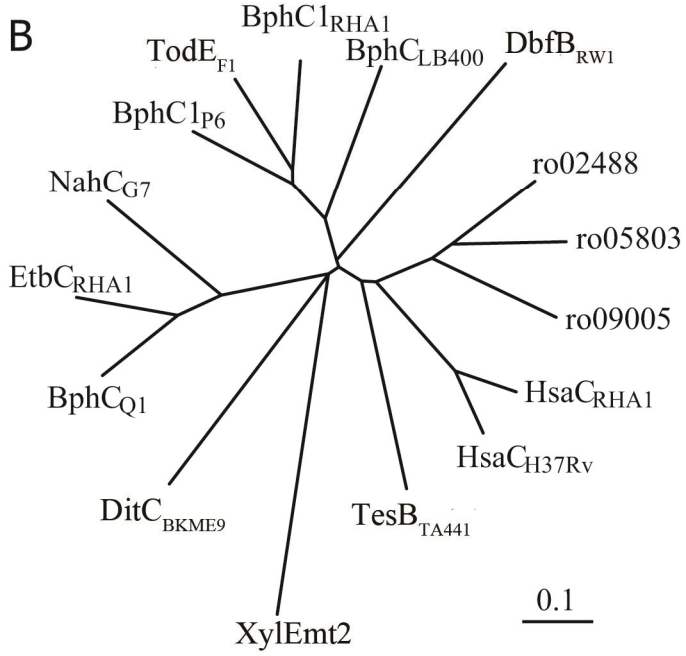
Wilbrink, M.H., Dijkhuizen, L., and van der Geize, R. FadD19 of *Rhodococcus rhodochrous* DSM43269: a steroid-CoA ligase essential for degradation of C24-branched sterol side chains. Submitted

Wilbrink, M.H., van der Geize, R., and Dijkhuizen, L. The *ltp3* and *ltp4* genes of *Rhodococcus rhodochrous* DSM43269 are essential for side chain degradation of C24-branched sterols. In preparation for submission

SUPPLEMENTAL DATA

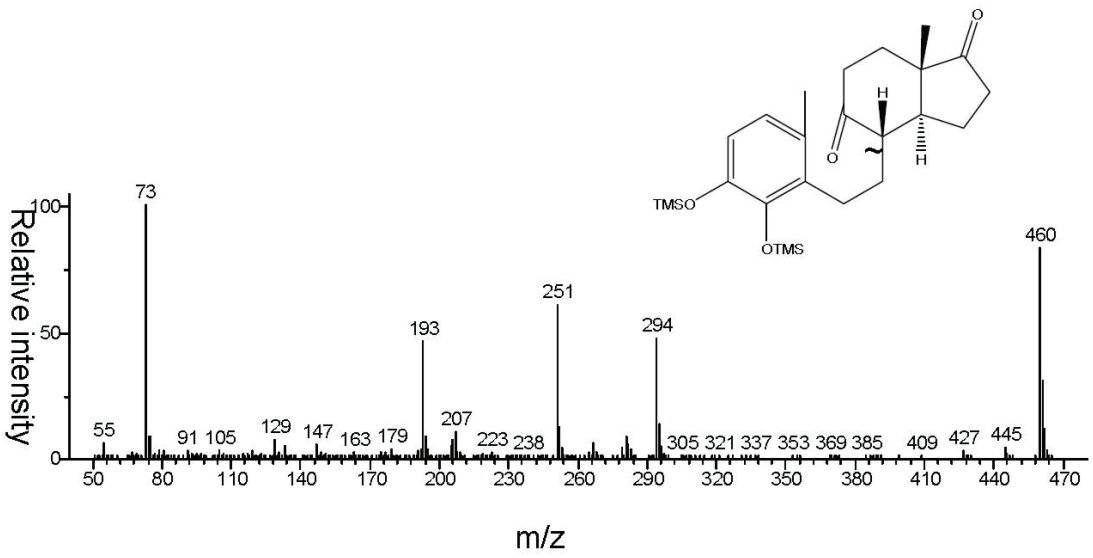
Supplemental data for **Chapter 2**





Supplemental Fig. S1. Phylogenetic analyses of KshA (A) and HsaC (B). KshA was analyzed by using the three catalytic subunits of phylogenetically diverse Rieske nonheme oxygenases (Ferraro *et al*., 2005; accession number in parentheses): BphAa_RHA1, biphenyl dioxygenase of RHA1 (GI:510285); EtbAa_RHA1, ethyl benzene dioxygenase of RHA1 (GI:35764426); CarAa_J3, carbazole dioxygenase of *Janthinobacterium sp*. J3 (GI:75765412); PDO_DB01, phthalate dioxygenase of *Burkholderia*

cepacia DB01 (GI:4128221); OMO_86, 2-oxoquinoline 8-monooxygenase of *Pseudomonas putida* 86 (GI:67463932); VanA_HR199, vanillate O-demethylase of *Pseudomonas* sp. HR199 (GI:3915234); and KshA_SQ1, 3-ketosteroid 9α-hydroxylase of *R. erythropolis SQ1* (GI:21309823). HsaC was analyzed by using type I extradiol dioxygenases that usually act on substrates containing two aromatic rings (Eltis and Bolin, 1996): BphC1_RHA1, 2,3-dihydroxybiphenyl dioxygenase (DHBD) of RHA1 (GI:510289); TodE_F1, 3-methylcatechol 2,3-dioxygenase of *P. putida* F1 (GI:151606); BphC1_P6, DHBD of *R. globerulus* P6 (GI:473117); BphC_LB400, DHBD of *B. xenovorans* LB400 (GI:1345621); DbfB, 2,2',3-trihydroxybiphenyl dioxygenase of *Sphingomonas wittichii* RW1 (GI:6226591); EtbC_RHA1, 3-ethylcatechol 2,3-dioxygenase of RHA1 (GI:1503984); BphC_Q1, DHBD of *Sphingomonas paucimobilis* Q1 (GI:115106); 17 NahC_G7, 1,2-dihydroxynaphthalene dioxygenase (DHBD) of *P. putida* G7 (GI:90567937); DitC_BKME9 dioxygenase of *Pseudomonas abietaniphila* BKME-9 (GI:4455074); and TesB_TA441, 3,4-DHSA dioxygenase of *C. testosteroni* TA441 (GI:10566463). XylE_mt2, catechol 2,3-dioxygenase of *P. putida* mt-2 (GI:139843) was used as an outgroup. Homologs of KshA and HsaC in RHA1 that are predicted to be involved in steroid degradation but whose specific function have not been experimentally demonstrated were identified by using Ro numbers (www.rhodococcus.ca). Sequences were aligned by using ClustalX (Jeanmougin *et al.*, 1998). Trees were visualized by using TreeView (Page, 1996).



Supplemental Fig. S2. Mass spectrum of trimethyl-silane (TMS)-derivatized 3,4-DHSA. The intensities of the ions were adjusted relative to that of the TMS peak at m/z = 73. The molecular ion (m/z = 460) corresponds to derivatized 3,4-DHSA (structure shown in Inset). The ion at m/z = 294 results from fragmentation between carbons 7 and 8, as depicted in the structure.

Supplemental Table S1. Annotation and expression of genes discussed in text.

Gene[a]	RHA1[b]	H37Rv[c]	ID[d]	Annotation of gene product	Best hit[e]	ID[f]	Expression ratio[g]
	ro00440	NA	NA	Transcriptional regulator	NA	NA	3.4*
prmA	ro00441	NA	NA	Propane monooxygenase, α-subunit	BAD03956	92	86
prmB	ro00442	NA	NA	Propane monooxygenase, reductase	BAD03957	63	7.8
prmC	ro00443	NA	NA	Propane monooxygenase, β-subunit	BAD03958	64	16.*
prmD	ro00444	NA	NA	Propane monooxygenase coupling protein	BAD03959	90	48*
	ro00445	NA	NA	Aminohydrolase 2	NA	NA	19*
	ro00446	NA	NA	Hypothetical protein	NA	NA	46*
	ro00447	NA	NA	NAD(H)-dependent Zn-binding alcohol dehydrogenase	1H2BA	43	9.6
	ro00448	Rv0440	50	Chaperonin GroEL	P40171	52	2.7
	ro00449	Rv1117	50	Flavoredoxin	NA	NA	20*
	ro00450	NA	NA	Glycolate oxidase, FAD-linked subunit	NA	NA	1.4
	ro00451	NA	NA	Hypothetical protein	NA	NA	5.1*
	ro00452	NA	NA	Transcriptional regulator, Fis family	NA	NA	46*
	ro00453	NA	NA	Hypothetical protein	NA	NA	9.6
	ro03459	Rv0044c	62	N5,N10-methylenetetrahydro-methanopterin reductase	AAD32732	38	2.4
fadA3	ro03460	Rv1074c	81	Acetyl-coA acetyltransferase	AAD22035	41	1.3
	ro03461	NA	NA	Feruloyl-coA synthetase	CAC18323	37	5.3
	ro03462	Rv1431	41	Membrane protein	NA	NA	4.4
	ro03463	Rv1432	66	Dehydrogenase	NA	NA	4.1
	ro03464	Rv1429	27	Regulator	NA	NA	6.4
choD	ro04305	Rv3409c	60	Cholesterol oxidase	P12676	19	1.1
kstR	ro04482	Rv3574	65	TetR-type transcriptional regulator	AAF19053	33	2.1
fadE34	ro04483	Rv3573c	57	Acyl-CoA dehydrogenase	NA	NA	2.5
fadE28	ro04484	Rv3544c	49	Short/branched chain acyl-CoA dehydrogenase	NA	NA	2.4
fadE29	ro04485	Rv3543c	73	Acyl-CoA dehydrogenase	BAD66691	28	1.5
	ro04486	Rv3542c	62	Actinomycete conserved hypothetical protein	NA	NA	1.0
	ro04487	Rv3541c	66	Actinomycete conserved hypothetical protein	NA	NA	2.0
ltp2	ro04488	Rv3540c	77	Branched-chain 3-ketoacyl-CoA thiolase	AAD21068	30	1.9
hsd4B	ro04531	Rv3538	62	2-Enoyl acyl-CoA hydratase	CAA55037	30[h]	3.2
kstD	ro04532	Rv3537	62	3-Ketosteroid-Δ1-dehydrogenase	AAL82579	40	3.3
hsaE	ro04533	Rv3536c	71	2-Hydroxypentadienoate hydratase	BAB97166	59	2.0
hsaG	ro04534	Rv3535c	85	Acetaldehyde dehydrogenase	BAB97164	61	3.2
hsaF	ro04535	Rv3534c	79	4-Hydroxy-2-oxovalerate aldolase	P51017	48	5.1*
	ro04537	Rv3527	27	Hypothetical protein	NA	NA	2.6
kshA	ro04538	Rv3526	59	Ketosteroid-9α-hydroxylase, oxygenase	AAL96829	57	4.6

hsaA	ro04539	Rv3570c	76	3-HSA hydroxylase, oxygenase	BAC67691	37	20*
hsaD	ro04540	Rv3569c	75	4,9-DSHA-hydrolase	BAC67693	31	9.2*
hsaC	ro04541	Rv3568c	81	3,4-DHSA dioxygenase	BAB15809	42	2.1
hsaB	ro04542	Rv3567c	70	3-HSA hydroxylase, reductase	BAC67692	16	9.3
	ro04586	Rv3519	50	Hypothetical protein	NA	NA	2.1*
	ro04587	NA	NA	Probable short-chain dehydrogenase	AAK47948	62	2.0*
	ro04588	Rv3518c	61	Cytochrome P450	CAC11139	34	1.0
fadE33	ro04591	Rv3564	42	Short/branched chain acyl-CoA dehydrogenases	BAD66690	33	1.9
fadE32	ro04592	Rv3563	55	Short/branched chain acyl-CoA dehydrogenases	BAD66685	27	1.1
fadE31	ro04593	Rv3562	65	Acyl-CoA dehydrogenase	BAD66691	34	2.0
fadD3	ro04595	Rv3561	60	Acyl-CoA ligase	BAC81695	44	3.5
fadE30	ro04596	Rv3560c	63	Acyl-CoA dehydrogenase	BAD66691	51	1.9
	ro04597	Rv3559c	72	Short chain dehydrogenase/reductase	BAD66692	60	0.70
	ro04598	Rv3557c	54	TetR-type transcriptional regulator	AAK50621	25	3.9*
fadA6	ro04599	Rv3556c	74	Acetyl-CoA acetyltransferase	BAD66686	63	1.9
	ro04649	Rv3553	71	2-Nitropropane dioxygenase	BAD66682	45	4.2
	ro04650	Rv3552	68	ATP-dependent CoA-transferase β-subunit	BAB15811	36	4.1*
	ro04651	Rv3551	74	ATP-dependent CoA-transferase α-subunit	BAB15810	42	4.7
echA20	ro04652	Rv3550	71	Enoyl-CoA hydratase	BAD66683	40	1.6
	ro04653	Rv3549c	62	Short-chain dehydrogenase/reductase	AAB08021	36	1.6
	ro04654	Rv3548c	71	Short-chain dehydrogenase/reductase	BAD66689	50	1.4
	ro04677	Rv3547	54	Hypothetical protein	BAB72053	36	1.3
fadA5	ro04678	Rv3546	74	Acetyl-CoA acetyltransferase	AAR83740	44	1.7
	ro04679	Rv3545c	62	Cytochrome P450	CAJ34365	27	11
ltp3	ro04683	Rv3523	79	SCPx related 3-ketoacyl-CoA thiolase	AAA40098	20	3.7
ltp4	ro04684	Rv3522	72	3-ketoacyl-CoA thiolase	NA	NA	1.3
	ro04685	Rv3521	62	Hypothetical protein	NA	NA	3.4
	ro04686	Rv3520c	72	FMN-dependent tetrahydro-methanopterin reductase	AAA92087	24	1.8
echA19	ro04688	Rv3516	73	Fatty acid-CoA hydratase	P31551	33	1.8
fadD19	ro04689	Rv3515c	64	Fatty acid-CoA ligase	AAB87139	38	12
fadD17	ro04691	Rv3506	56	Fatty acid-CoA ligase	Q4LDG0	20	3.6
fadE27	ro04692	Rv3505	54	Acyl-CoA dehydrogenase	P16219	24	1.7
fadE26	ro04693	Rv3504	77	Acyl-CoA dehydrogenase	P71539	25	9.5
fdxD	ro04694	Rv3503c	34	Ferredoxin	AAK81833	29	ND
hsd4A	ro04695	Rv3502c	59	17β-hydroxysteroid dehydrogenase	BAD66689	36[i]	7.2
yrbE4A	ro04696	Rv3501c	72	ABC transporter permease subunit	AAT51759	55	3.6
yrbE4B	ro04697	Rv3500c	66	ABC transporter permease subunit	AAT51760	49	2.7
mce4A	ro04698	Rv3499c	41	MCE family protein	CAA50257	32	2.3
mce4B	ro04699	Rv3498c	46	MCE family protein	NA	NA	2.2
mce4C	ro04700	Rv3497c	41	MCE family protein	NA	NA	2.5
mce4D	ro04701	Rv3496c	43	MCE family protein	NA	NA	1.7
mce4E	ro04702	Rv3495c	35	MCE family protein	NA	NA	3.0
mce4F	ro04703	Rv3494c	37	MCE family protein	NA	NA	0.76

	ro04704	Rv3493c	31	MCE associated protein	NA	NA	0.69
	ro04705	Rv3492c	26	MCE associated protein	NA	NA	0.78
kshB	ro05833	Rv3571	54	KSH, reductase	AAL96830	69	1.5
	ro06687	Rv1400c	36	Carboxylesterase	1EVQA	41	6.8
	ro06688	NA	NA	Carboxylesterase	1M33A	23	6.7*
	ro06689	NA	NA	Hypothetical protein	NA	NA	15*
	ro06690	NA	NA	NADH pyrophosphatase	1VK6A	27	6.9
	ro06691	NA	NA	Transcriptional regulator	AAK73166	43	4.9
	ro06692	NA	NA	Fatty acid-CoA ligase	1V26A	36	2.2*
	ro06693	NA	NA	5-Valerolactone hydrolase	BAC22650	68	8.1*
	ro06694	NA	NA	2-Nitropropane dioxygenase	BAD66682	27	6.0
	ro06695	NA	NA	Enoyl-CoA hydratase	ABC33869	32	1.8*
	ro06696	NA	NA	Hypothetical protein	NA	NA	1.6
	ro06697	NA	NA	Acyl-coA dehydrogenase	NA	NA	ND
	ro06698	NA	NA	Cyclohexanone monooxygenase	AAG01290	59	5.3
	ro08053	NA	NA	Hypothetical protein	NA	NA	3.9
bphB1	ro08054	NA	NA	2,3-Dihydroxy-1-phenylcyclo-hexa-4,6-diene dehydrogenase	BAA06873	100	2.8
bphC1	ro08055	NA	NA	Biphenyl-2,3-diol 1,2-dioxygenase	BAA06872	100	3.5
bphA4	ro08057	NA	NA	Biphenyl dioxygenase, reductase subunit	BAA06871	100	6.1
bphA2	ro08059	NA	NA	Biphenyl 2,3-dioxygenase, β-subunit	BAA06869	100	4.5
bphA1	ro08060	NA	NA	Biphenyl 2,3-dioxygenase, α-subunit	BAA06868	100	3.5
cpmA	ro08137	NA	NA	Cyclopentanone 1,2-monooxygenase	CAD10798	49	2.0
	ro08138	NA	NA	Acyl-CoA dehydrogenase	AAK18172	32	8.4
cpnC	ro08142	NA	NA	5-Valerolactone hydrolase	BAC22650	64	4.6
	ro08143	NA	NA	2-Nitropropane dioxygenase	BAD66682	28	10
	ro08145	NA	NA	Hypothetical protein	NA	NA	5.5
	ro08147	NA	NA	Transposase fragment	NA	NA	4.2
bphB2	ro10126	NA	NA	Alkylbenzene dihydrodiol dehydrogenase	AAR90124	100	3.8
chnE	ro10127	Rv0234c	39	6-Oxohexanoate dehydrogenase	AAN37492	64	3.4
etbC	ro10135	NA	NA	Biphenyl-2,3-diol 1,2-dioxygenase	BAC92714	100	3.7
bphD1	ro10136	NA	NA	2-Hydroxy-6-oxo-6-phenyl-hexa-2,4-dienoate hydrolase	BAC92715	100	3.8
ebdA1	ro10143	NA	NA	Ethylbenzene dioxygenase, α-subunit	BAC92718	100	3.3
ebdA2	ro10144	NA	NA	Ethylbenzene dioxygenase, β-subunit	BAC92719	100	3.2
	ro10158	NA	NA	Hypothetical protein	NA	NA	4.3
acxB	ro10159	NA	NA	Acetone carboxylase, α-subunit	AAL17711	24	5.6
acxA	ro10160	NA	NA	Acetone carboxylase, β-subunit	AAL17710	28	3.0
acxC	ro10161	NA	NA	Acetone carboxylase, γ-subunit	NA	NA	6.4
	ro10162	NA	NA	Acetone carboxylase subunit	NA	NA	2.3

[a] Gene name assigned based on current study

[b] RHA1 gene identification

[c] *M. tuberculosis* H37Rv gene identification for reciprocal best hit to RHA1 gene

[d] Percent amino acid sequence identity of the RHA1 and H37Rv orthologues based on full sequence alignment

[e] Accession number of functionally characterized best hit in NCBI database

[f] Percent amino acid sequence identity of the RHA1 enzyme and its experimentally characterized best hit based on full sequence alignment

[g] Average normalized expression ratio (cholesterol/pyruvate); asterisk indicates $p < 0.05$ for Student's *t*-test

[h] Identity with central domain a.a. 324-596

[i] Identity with N-terminal domain a.a. 1-323

Supplemental Table S2. Annotation of RHA1, H37Rv and BCG genes assigned to cholesterol pathway.

Gene[a]	RHA1[b]	H37Rv[c]	BCG[d]	ID[e]	Annotation of gene product	Best hit[f]	ID[g]
mce4F	Ro04703	Rv3494c	Bcg3558c	37	MCE family protein	NA	NA
mce4E	Ro04702	Rv3495c	Bcg3559c	35	MCE family protein	NA	NA
mce4D	Ro04701	Rv3496c	Bcg3560c	43	MCE family protein	NA	NA
mce4C	Ro04700	Rv3497c	Bcg3561c	41	MCE family protein	NA	NA
mce4B	Ro04699	Rv3498c	Bcg3562c	46	MCE family protein	NA	NA
mce4A	Ro04698	Rv3499c	Bcg3563c	41	MCE family protein	CAA50257	32
supB	Ro04697	Rv3500c	Bcg3564c	66	Sterol uptake permease subunit (ABC transporter)	AAT51760	49
supA	Ro04696	Rv3501c	Bcg3565c	72	Sterol uptake permease subunit (ABC transporter)	AAT51759	55
choD	Ro04305	Rv3409c	Bcg3479c	60	cholesterol oxidase	P12676	19
hsd4A	Ro04695	Rv3502c	Bcg3566c	59	17β-hydroxysteroid dehydrogenase	BAD66689	36
hsd4B	Ro04531	Rv3538	Bcg3602	62	2-Enoyl acyl-CoA hydratase	CAA55037	30
kshA	Ro04538	Rv3526	Bcg3590	59	Ketosteroid-9α-hydroxylase, oxygenase	AAL96829	57
kshB	Ro05833	Rv3571	Bcg3636	54	Ketosteroid-9α-hydroxylase, reductase	AAL96830	69
kstD	Ro04532	Rv3537	Bcg3601	62	3-ketosteroid-Δ1-dehydrogenase	AAL82579	40
hsaB	Ro04542	Rv3567c	Bcg3632c	70	3-HSA hydroxylase, reductase	BAC67692	16
hsaC	Ro04541	Rv3568c	Bcg3633c	81	3,4-DHSA dioxygenase	BAB15809	42
hsaD	Ro04540	Rv3569c	Bcg3634c	75	4,9-DSHA hydrolase	BAC67693	31
hsaA	Ro04539	Rv3570c	Bcg3635c	76	3-HSA hydroxylase, oxygenase	BAC67691	37
hsaF	Ro04535	Rv3534c	Bcg3598c	79	4-hydroxy-2-oxovalerate aldolase	P51017	48
hsaG	Ro04534	Rv3535c	Bcg3599c	85	Acetaldehyde dehydrogenase	BAB97164	61
hsaE	Ro04533	Rv3536c	Bcg3600c	71	2-hydroxypentadienoate hydratase	BAB97166	59
fadE26	Ro04693	Rv3504	Bcg3568	77	acyl-CoA dehydrogenase	P71539	25
fadE27	Ro04692	Rv3505	Bcg3569	54	acyl-CoA dehydrogenase	P16219	24
fadD17	Ro04691	Rv3506	Bcg3570	56	Fatty acid-CoA synthetase	Q4LDG0	20
fadD19	Ro04689	Rv3515c	Bcg3578c	64	Fatty acid-CoA ligase	AAB87139	38
echA19	Ro04688	Rv3516	Bcg3579	73	Fatty acid-CoA hydratase	P31551	33
ltp4	Ro04684	Rv3522	Bcg3586	72	3-ketoacyl-CoA thiolase	NA	NA
ltp3	Ro04683	Rv3523	Bcg3587	79	SCPx related 3-ketoacyl-CoA thiolase	AAA40098	20
	Ro06698	NA	NA	NA	Cyclohexanone monooxygenase	AAG01290	59
	Ro06693	NA	NA	NA	5-Valerolactone hydrolase	BAC22650	68

[a]Name assigned based on current study. Identification numbers for the corresponding [b]RHA1 gene and the reciprocal best hit in [c]*M. tuberculosis* H37Rv and [d]*M. bovis* BCG. [e]Percent amino acid sequence identity of the RHA1 and H37Rv and BCG orthologues based on full sequence alignment. Nucleotide sequence identity between H37Rv and BCG genes is >98%. [f]Accession number of functionally characterized best hit in NCBI database. [g]Percent amino acid sequence identity of the

RHA1 enzyme and its experimentally characterized best hit based on full sequence alignment. NA, not available, either no homologous gene in H37Rv or BCG, or no functionally characterized homologue available.

Supplemental Table S3. Oligonucleotides used.

Name	Use	Sequence
hsaC-for3	qRT-PCR of *hsaC*	TCCGCAAGAAGGTCAAGATGT
hsaC-rev3	qRT-PCR of *hsaC*	TCTTCATGTAGAACGACAGCATCA
hsaC-TP	Taqman probe *hsaC*	FAM-CCGTCACGTCAACG-MGB
hsaC-for1	*hsaC* replacement	TCTGGCCACCGATTTCCTTCTGGACGGGGGTAAGTGAT GATTCCGGGGATCCGTCGACC
hsaC-rev1	*hsaC* replacement	CCTCGGCCGCGACGGCCCCGTCCCCGGTCACCTCACTCA TGTAGGCTGGAGCTGCTTC
hsa-for2	verify deletion of *hsaC* (internal)	GTATGGGCACAAGTTCGTCA
hsa-rev2	verify deletion of *hsaC* (internal)	GGAACGCGAGACTGTGATG
hsa-for5	verify deletion of *hsaC* (external)	AGTCGATCAGCCGGGATAC
hsa-rev6	verify deletion of *hsaC* (external)	GTCGCTCCTCCCTGATCTC
ro04698-F	*mce4* cluster deletion	ATACTTCGTCTTCTCGGTCTACATCC
ro04698-R (*Spe*I)[a]	*mce4* cluster deletion	ATACTAGTCTTCCCACCCGAATCAGTCATCT
ro04703-F (*Spe*I)[a]	*mce4* cluster deletion	ATAACTAGTTGAGTTCCAGCACAGACGAGCG
ro04703-R (*Hind*III)	*mce4* cluster deletion	ATAAGCTTGCGCGATGGTTTCCTGCACGAGG
SupA-F	*yrbE4AB* deletion	GGCTGACCGCGATGTGAACGTGGA
SupA-R(*Spe*I)[b]	*yrbE4AB* deletion	AACTAGTCGCACGCAACGGCACCTCGAGGA
SupB-F(*Spe*I)[b]	*yrbE4AB* deletion	AACTAGTGGTTTCAACGCCTCGGGCGGTCCG
SupB-R	*yrbE4AB* deletion	GGGCTTGAGCCCGGACACGTCGGT
Mce4A-F	Verify deletion of *mce4* cluster	ATGGAACCCGACGCCAAGGTCAA
Mce4A-R	Verify deletion of *mce4* cluster	CTGTTCGGCGAGGGGGATCACAT
Sup-F	Verify deletion of *yrbE4AB*	TATGGTAGACCTCCTCGAGGTGCCG
Sup-R	Verify deletion of *yrbE4AB*	CTATCCCGAAAGGTTGAAGTTGCCG
Hcmt-F	amplification of *hsaC* H37Rv	CGACTAGCATATGAGCATCCGGTCGC
Hcmt-R	amplification of *hsaC* H37Rv	CGGGATCCCTGAGCCGACATCGTTTG
Hdmt-F	amplification of *hsaD* H37Rv	CGACGTACATATGACAGCTACCGAGGAATTG
Hdmt-R	amplification of *hsaD* H37Rv	CAGGATCCTCATCTGCCACCTCCCAG

Introduced restriction sites are underlined. [a,b] The *Spe*I restriction sites were designed to enable in-frame linkage of the *ro04698*[a] or *ro04696*[b] start codon with the *ro04703*[a] or *ro04697*[b] stop codon, effectively replacing the 6 *mce4* genes[a] or the 2 *yrbE4AB* genes[b] with a 30 nt[a] or 186 nt[b] remnant, respectively.

References

Ferraro, D.J., Gakhar, L., and Ramaswamy, S. (2005) Rieske business: structure-function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Commun.* **338:** 175-190

Eltis, L.D., and Bolin, J.T. (1996) Evolutionary relationships among extradiol dioxygenases. *J. Bacteriol.* **178:** 5930-5937

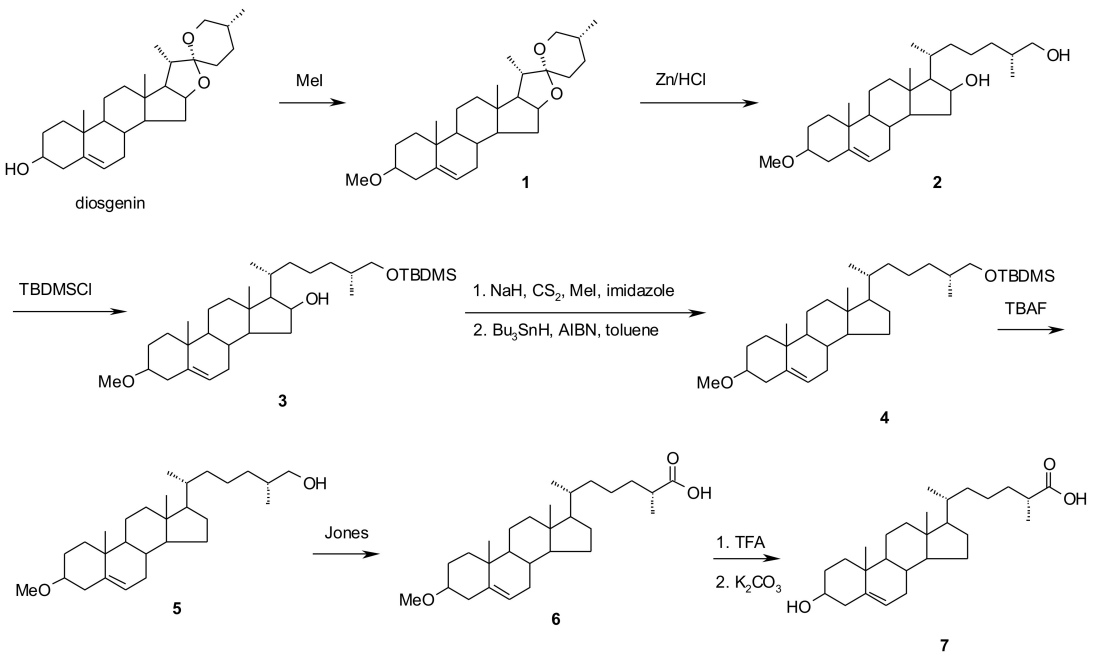
Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. (1998) Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23:** 403-405

Page, R.D.M. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comp. Appl. Biosci.* **12:** 357-358

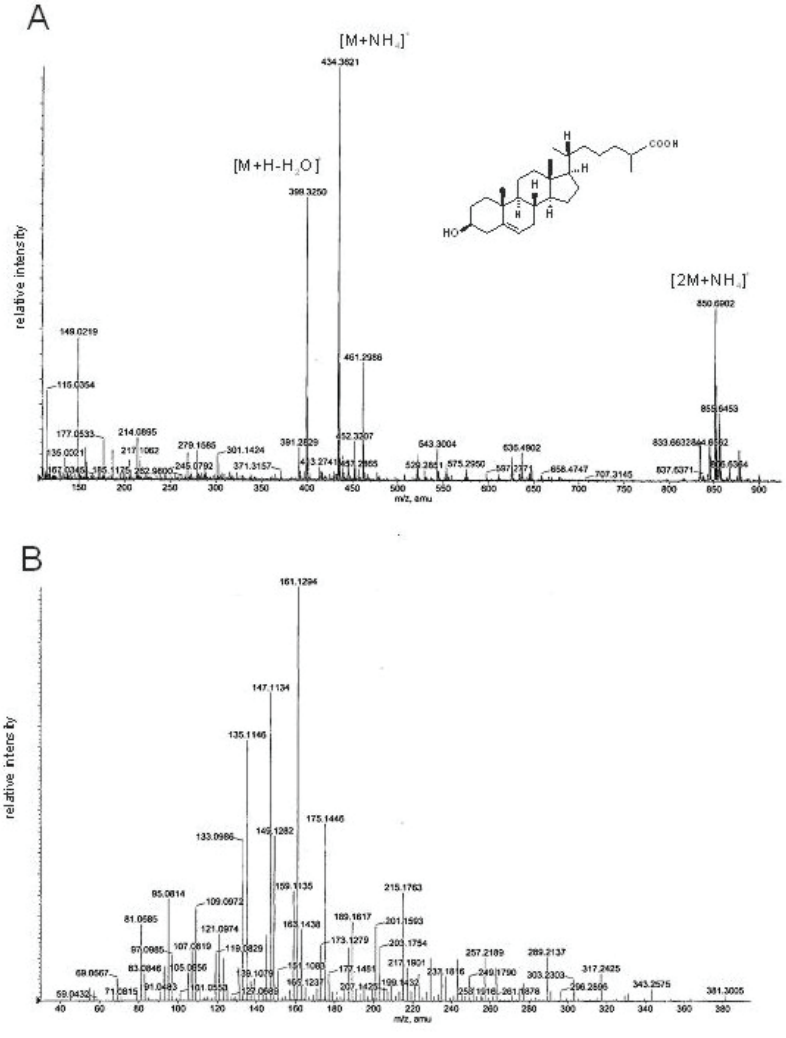
Supplemental data for **Chapter 3**

```
                                  GxxT
CYP125_RHA1      235 LSPEEFGFFVILLAVAGNETTRNAITHGMMAFLDHPD
CYP125_DSM43269  235 LAPEEFGFFFIVLAVAGNETTRNAITHGMAAFLDNPE
CYP125_H37Rv     253 LSDDEFGFFVVMLAVAGNETTRNSITQGMMAFAEHPD
CYP11A1          272 MLLEDVKANITEMLAGGVNTTSMTLQWHLYEMARSLN
CYP27A1h         287 LSPREAMGSLPELLMAGVDTTSNTLTWALYHLSKDPE
CYP3A4h          290 LSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPD
CYP46A1h         287 QDDEGLLDNFVTFFIAGHETSANHLAFTVMELSRQPE

                        ExxR    V295 R300                                   PxxF   R343
CYP125_RHA1      282 KTTADEIVRWATPVNSFQRTALEDTELGGVQIKKGQRVVMLYGSANFDEDAFENPEKFDIMR
CYP125_DSM43269  282 KTAADEIIRWATPVTSFQRTALEDTELGGQTIRKGERVVMLYASANNDEEVFENPREFDILR
CYP125_H37Rv     300 ETAADEIVRWATPVTAFQRTALRDYELSGVQIKKGQRVVMFYRSANFDEEVFQDPFTFNILR
CYP11A1          339 KASIKETLRLHPISVTLQRYPESDLVLQDYLIPAKTLVQVAIYAMGRDPAFFSSPDKFDPTR
CYP27A1h         354 KAVLKETLRLYPVVPTNSRIIEKEIEVDGFLFPKNTQFVFCHYVVSRDPTAFSEPESFQPHR
CYP3A4h          357 DMVVNETLRLFPIAMRLERVCKKDVEINGMFIPKGVVVMIPSYALHRDPKYWTEPEKFLPER
CYP46A1h         354 SQVLKESLRLYPPAWGTFRLLEEETLIDGVRVPGNTPLLFSTYVMGRMDTYFEDPLTFNPDR

                      FxxxGxxxCxG
CYP125_RHA1      349 VGFGGTGAHFCLGANLARLEIDL
CYP125_DSM43269  349 LAFGGTGAHYCLGANLARMEIDL
CYP125_H37Rv     367 VGFGGTGAHYCIGANLARMTINL
CYP11A1          414 LGF-GWGVRQCVGRRIAELEMTL
CYP27A1h         434 VPF-GYGVRACLGRRIAELEMQL
CYP3A4h          433 TPF-GSGPRNCIGMRFALMNMKL
CYP46A1h         428 FPF-SLGHRSCIGQQFAQMEVKV
```

Supplemental Fig. S1. Partial amino acid sequence alignment of actinobacterial and mammalian cytochrome P450 proteins: *R. jostii* RHA1 CYP125 (accession number ABG96465), *R. rhodochrous* DSM43269 CYP125 (FJ824698), *M. tuberculosis* H37Rv CYP125 (NP_218062), bovine CYP11A1 (NP_788817), human CYP27A1 (NP_000775), human CYP3A4 (P08684) and human CYP46A1 (NP_006659). The alignment was generated using ClustalW, MEGA version 3.1 (Kumar *et al*., 2004). Specific residues discussed in the text in bold are indicated (V295, R300 and R343 of $Cyp125_{RHA1}$). Structure homology modeling of the CYP11A1:cholesterol complex revealed that the conserved arginine (R357, corresponding to R300 of $CYP125_{RHA1}$) points towards the docked cholesterol and is believed to be involved in catalysis (Storbeck *et al.,* 2007). Moreover, this residue was shown to be one of the active site residues in CYP46A1 (Mast *et al*., 2008) and has been recently identified in P450s (e.g. CYP3A4) as a heme-interacting residue involved in substrate regio-selectivity and specificity (Seifert and Pleiss, 2009). For its part, the conserved arginine of CYP27A1 (R415, corresponding to R343 of $CYP125_{RHA1}$) was shown to be connected via a tripartite salt-bridge with the ExxR motif, forming an ERR-triad as part of the active site cavity (Prosser *et al.,* 2006; Masuda *et al.,* 2007). Finally, in CYP27A1, V367 (corresponding to V295 of $CYP125_{RHA1}$) is an active site residue and appears to be crucial for regioselectivity of hydroxylation of both cholesterol and 5β-cholestane 3α,7α,12α-triol (Mast *et al.,* 2006). Conserved motifs of the P450s super-family (ExxR and PxxF), the oxygen-binding domain (GxxT) and the heme binding domain (F[x]xxGxxxCxG) (Mestres, 2005; Huang *et al*., 2008) are also shown.

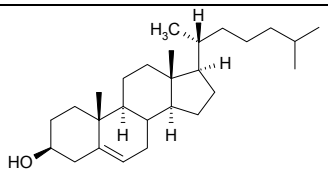
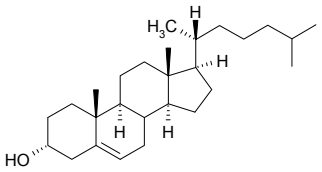
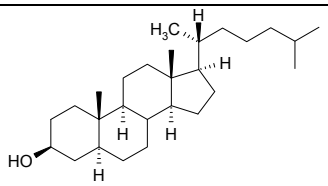
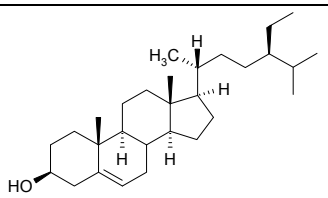
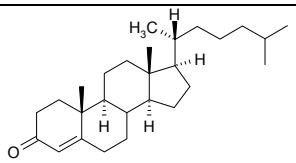
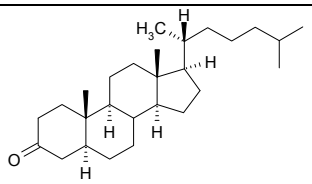
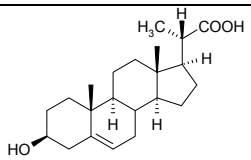
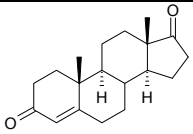
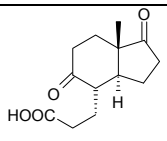


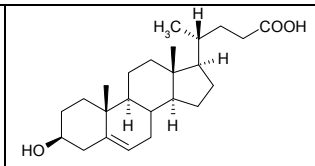
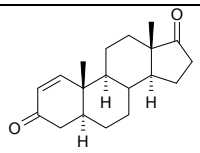
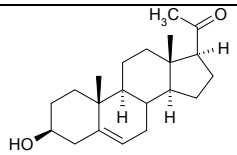
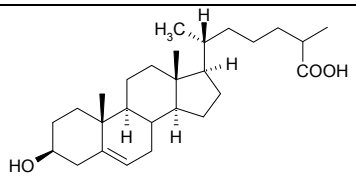
Supplemental Fig. S2. Scheme of the chemical synthesis of 5-cholestene-26-oic-acid-3β-ol from diosgenin.



Supplemental Fig. S3. (A) MS spectrum with fragmentation pattern of chemically synthesized 5-cholestene-26-oic acid-3β-ol. (B) MS-MS spectrum with fragmentation pattern of ion peak with m/z 399.

Supplemental Table S1. Growth of wild type *Rhodococcus jostii* strain RHA1 and mutant strain RHA1Δ*cyp125* on different sterols and related organic substrates. Growth is indicated with (+), whereas no growth is indicated with (-).

Steroid name	Steroid formula	RHA1	RHA1Δ*cyp125*
cholesterol (5-cholestene-3β-ol)		+	-
Epicholesterol (5-cholestene-3α-ol)		+	-
5α-cholestanol (5α-cholestane-3β-ol)		+	-
β-sitosterol* (5-cholestene-24β-ethyl-3β-ol)		+	-
Cholestenone (4-cholestene-3-one)		+	+
cholestanone (5α-cholestane-3-one)		+	+
23,24-bisnor-5-cholenic acid-3β-ol		+	+
4-androstene-3,17-dione		+	+
9,17-dioxo-1,2,3,4,10,19-hexanorandrostane-5-oic acid		+	+

5-cholenic acid-3β-ol		+	+
1-5α-androstene-3,17-dione		+	+
5-pregnene-3β-ol-20-one		+	+
5-cholestene-26-oic acid-3β-ol		+	+

* Mixture of β-sitosterol (>75 %) with β-sitostanol (10-14 %) and campesterol (6-9 %) was used.

Supplemental Table S2. Strains and plasmids used in this study.

Strain	**Characteristics**	**Reference**
Rhodococcus jostii RHA1	Wild type strain, PCB degrader	(Masai *et al*., 1995)
Rhodococcus rhodochrous DSM43269	Wild type strain; potent sterol degrader; identical to strain IFO3338	DSMZ culture collection
Rhodococcus rhodochrous RG32	Mutant strain of DSM43269 capable of selective sterol side chain degradation; carries 5-fold unmarked *kshA* gene deletion	Wilbrink *et al.,* submitted
E. coli DH5α	General host for cloning	Bethesda Res. Laboratories
E. coli S17-1	Host strain for conjugal mobilization of pK18mobsacB derivatives to *Rhodococcus* strains	DSMZ culture collection
Rhodococcus jostii RHA1Δ*cyp125*	*cyp125* gene deletion mutant strain of RHA1; blocked side chain degradation	This study
Rhodococcus rhodochrous RG32Ω*cyp125*	*cyp125* gene disruption mutant strain of RG32; blocked side chain degradation	This study
Plasmid		
pK18*mobsacB*	Conjugative plasmid for gene mutagenesis in *Rhodococcus*; *aphII sacB oriT* (RP4) *lacZ*	(van der Geize *et al*., 2001)
pBlueScript KS(II)	General *E. coli* cloning vector; *bla lacZ*	Stratagene
pTip-QC1	*Rhodococcus* expression vector; Chlr *bla PtipA repAB* (pRE2895)	(Nakashima and Tamura, 2004)
pRESQ	*E.coli-Rhodococcus* shuttle-vector; *aphII, lacZ-ccdB rep* (pMVS301)	(van der Geize *et al*., 2002b)
pBs-Apra-ori	Conjugative pBlueScript KS(II) derivative; *aacIV*, *oriT* (RP4) *lacZ*	(van der Geize *et al.*, 2008a)
pRRE1	*E. coli- R. rhodochrous* shuttle-vector	This study
pDEL*cyp125*$_{RHA1}$	pK18*mobsacB*-derived mutagenic plasmid for *cyp125*$_{RHA1}$ disruption in RHA1	This study
pTip-QC1*cyp125*$_{RHA1}$	*cyp125*$_{RHA1}$ expression plasmid used for CYP125$_{RHA1}$ production and functional complementation of RHA1□*cyp125*	This study

pRESQ4679	Genomic library clone DSM43269 carrying c*yp125*$_{DSM43269}$	This study
pΩ*cyp125*	pK18*mobsacB* derived mutagenic plasmid for *cyp125* gene disruption in RG32	This study
pCOMP*cyp125* $_{DSM43269}$	pRRE1-derived plasmid carrying *cyp125*$_{DSM43269}$; functional complementation of RG32Ωcyp125	This study

References

Huang, S., Sun, D., and Brattsten, L.B. (2008) Novel cytochrome P450s, CYP6BB1 and CYP6P10, from the salt marsh mosquito *Aedes sollicitans* (Walker) (Diptera: Culicidae). *Arch. Insect. Biochem. Physiol.* **67:** 139-154

Kumar, S., Tamura, K., and Nei, M. (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinformat.* **5:** 150-163

Mast, N., Murtazina, D., Liu, H., Graham, S.E., Bjorkhem, I., Halpert, J.R., *et al*. (2006) Distinct binding of cholesterol and 5beta-cholestane-3alpha,7alpha,12alpha-triol to cytochrome P450 27A1: evidence from modeling and site-directed mutagenesis studies. *Biochemistry* **45:** 4396-4404

Mast, N., White, M.A., Bjorkhem, I., Johnson, E.F., Stout, C.D., and Pikuleva, I.A. (2008) Crystal structures of substrate-bound and substrate-free cytochrome P450 46A1, the principal cholesterol hydroxylase in the brain. *Proc. Natl. Acad. Sci. USA* **105:** 9546-9551

Mestres, J. (2005) Structure conservation in cytochromes P450. *Proteins* **58:** 596-609

Prosser, D.E., Guo, Y., Jia, Z., and Jones, G. (2006) Structural motif-based homology modeling of CYP27A1 and site-directed mutational analyses affecting vitamin D hydroxylation. *Biophys. J.* **90:** 3389-3409

Seifert, A., and Pleiss, J. (2009) Identification of selectivity-determining residues in cytochrome P450 monooxygenases: A systematic analysis of the substrate recognition site 5. *Protein* **74:** 1028-103

Supplemental data for **Chapter 4**

Supplemental Table S1. Bacterial strains and plasmids used in this study.

Strain	Description	Reference
E. coli BL21(DE3)	Host for expression of T7 based plasmids	Novagen
E. coli DH5α	General host for cloning purposes	Bethesda Res. Laboratories
E. coli S17-1	Strain for conjugal mobilization of pK18*mobsac*B derived plasmids to *Rhodococcus* strains	DSMZ culture collection
Rhodococcus jostii RHA1	Wild type strain; PCB degrader	(Masai *et al*., 1995)
Rhodococcus rhodochrous DSM43269	Wild type strain; potent sterol degrader; identical to strain IFO3338	DSMZ culture collection
Rhodococcus rhodochrous RG32	Mutant strain of DSM43269 capable of selective sterol side chain degradation; carries 5-fold unmarked *kshA* gene deletion	Petrusma *et al.,* in preparation
Rhodococcus jostii MW1	Deletion mutant of RHA1 lacking *ro04683-ro04694*	This study
RG32Δ*ro04690*$_{DSM43269}$	Gene deletion mutant of *ro04690*$_{DSM43269}$ in *R. rhodochrous* strain RG32	This study
RG32Δ*ro04690*$_{DSM43269}$ Δ*fadE26*	Double deletion mutant of *ro04690*$_{DSM43269}$ and *fadE26* in *R. rhodochrous* strain RG32	This study
RG32Δ*ro04690*$_{DSM432690}$ Δ*fadE26* Ω*fadE27*	Gene inactivation mutant of *fadE27* in RG32Δ*ro04690*$_{DSM43269}$Δ*fadE26*	This study
RG32Δ*fadD19*	Gene deletion mutant of *fadD19* in *R. rhodochrous* strain RG32	This study
RG32 Ω*fadD17*	Gene inactivation mutant of *fadD17* in *R. rhodochrous* strain RG32	This study
RG32Δ*fadD19* Ω*fadD17*	Gene inactivation mutant of *fadD17* in RG32Δ*fadD19*	This study
Plasmid	**Description**	**Reference**
pK18*mobsacB*	Conjugative plasmid for gene mutagenesis in *Rhodococcus*; *aphII sacB oriT* (RP4) *lacZ*	(van der Geize *et al*., 2001)
pBlueScript(II)KS	General *E. coli* cloning vector; *bla lacZ*	Stratagene
pRESQ	*E. coli-Rhodococcus* shuttle-vector; *aphII, lacZ-ccdB rep* (pMVS301)	(van der Geize *et al*., 2002b)
pBs-Pkan	pBlueScript(II)KS containing *aphII* promoter region	(van der Geize *et al*., 2008a)
pET15b	T7 promotor based expression plasmid; *bla lacI*	Novagen
pET15b*fadD19*$_{DSM43269}$	pET15b containing *fadD19*$_{DSM43269}$	This study
pDELBox	pK18*mobsacB*-derived mutagenic plasmid for deletion of *ro04683-ro04694* in RHA1	This study
pDEL*fadD19*$_{DSM43269}$	pK18*mobsacB*-derived mutagenic plasmid for deletion of *fadD19*$_{DSM43269}$ in RG32	This study
pDEL*fadE26*$_{DSM43269}$	pK18*mobsacB*-derived mutagenic plasmid for deletion of *fadE26*$_{DSM43269}$ in RG32	This study
pDEL*ro04690*$_{DSM43269}$	pK18*mobsacB*-derived mutagenic plasmid for deletion of *ro04690*$_{DSM43269}$ in RG32	This study
pΩ*fadE27*$_{DSM43269}$	pK18*mobsacB*-derived mutagenic plasmid for disruption of *fadE27*$_{DSM43269}$ in RG32	This study
pΩ*fadD17*$_{DSM43269}$	pK18*mobsacB*-derived mutagenic plasmid for disruption of *fadD17*$_{DSM43269}$ in RG32	This study
pRESQ4690	Genomic library clone DSM43269 carrying *ro04690*$_{DSM43269}$	This study
pRESQ4693	Genomic library clone DSM43269 carrying *fadE26*$_{DSM43269}$	This study
pCOMP*fadD19*$_{DSM43269}$	*fadD19*$_{DSM43269}$ in pRESQ behind the *aphII* promotor; functional complementation	This study

Supplemental Table S2. List of primers used in this study. Introduced restriction sites are underlined.

Nr	PCR amplicon	Size (bp)	Oligonucleotide sequence (nr)
1	Upstream region *ro04683* of RHA1 (deletion construct MW1)	1,088	BoxUP-F GCCGTAGACGACGAAGACCTGA BoxUP-R CAGATCTCGAGTGGCGAGACGAGGAGAAG
2	Downstream region *ro04694* of RHA1 (deletion construct MW1)	1,178	BoxDOWN-F CAGATCTGGAGAAGCAGAGCGCCGACTAC BoxDOWN-R CAGGCCTTGCGCAGTGTCTTCA
3	*ro04683-ro04694* genes (confirmation deletion mutant MW1)	767	BoxContr-F ACCTTGCCGTCCACGTTCACA BoxContr-R TCATCTGCAGGCGGTCGAAGT
4	Internal fragment *ro04690* homologue (screening genomic library)	361	ro04690^{DEG}-F CACACCGG(C/T)GAG(A/G)T(C/G)GCG(A/T)C(C/G)ATG ro04690^{DEG}-R TCTGCAGCGGCAT(C/G)(C/G)(C/G)(A/C)AGCGG
5	Internal Fragment *fadE26* homologue (screening genomic library)	670	fadE26^{DEG}-F (C/G)(A/G)TCAACGG(A/C)CAGAAGATGTGG fadE26^{DEG}-R A(C/T)(C/T)TCGTTGGT(G/T)CCGCCGCCGAA
6	*fadD19*$_{DSM43269}$ gene (heterologous expression in pET15b)	1,679	fadD19exp-F TCATATGGCCCTAAACATCGCAGACC fadD19exp-R TAGATCTCTATCCCGTCGCCGCGGCGCCG
7	*fadD19*$_{DSM43269}$ gene (confirmation Δ*fadD19* mutant)	527 (wt: 1278)	fadD19Contr-F GTGAAGACCGTCGTGGTGGTAG fadD19Contr-R GTTGCCGTCGCCGTTGTTGTAG
8	*fadD19*$_{43269\,gene}$ (complementation Δ*fadD19* mutant)	1,673	COMPfadD19-F GGCCCTAAACATCGCAGACCTC COMPfadD19-R CGGTACCCTATCCCGTCGCCGCGGCGCCG
9	Internal fragment *fadD17*$_{DSM3269}$ (disruption construct)	826	fadD17-F GACATCGCGCTCGCCGATTGCC fadD17-R TGTAGTAGCCGTCGAACAGACC
10	*fadD17*$_{DSM43269}$ (confirmation *fadD17* disruption)	1,027	pK18Contr-F AATGCAGCTGGCACGACAGGTT fadD17Contr-R CGAGCATCCTGCGCGAACTCGG
11	Internal fragment *fadE27*$_{DSM3269}$ (disruption construct)	753	fadE27-F GCAGCATCGTCGCGGACATC fadE27-R GGAGCCGAGCAGGAACTCGT
12	*fadE27*$_{DSM43269}$ (confirmation *fadE27* disruption)	1,510	fadE27Contr-F GCAGCATCGTCGCGGACATC pK18Contr-R CTGCGTGCAATCCATCTTGTTC
13	Upstream region *ro04690*$_{DSM43269}$ (deletion construct)	1,247	ro04690UP-F AGCGCCGACGACATCTACATCC ro04690UP-R TCATATGCGTGAATCCGAAGATCGGATAC
14	Downstream region *ro04690*$_{DSM43269}$ (deletion construct)	1,236	ro04690DOWN-F TCATATGGAGATCATGGCCGAACTCGTC ro04690DOWN-R CTGCAGGATCACGGCAACGAC
15	*ro04690*$_{DSM43269}$ gene (confirmation Δ*ro04690* mutant)	129 (wt: 1,113)	ro04690Contr-F ATGCAGACCGGCCTCAGCAAGA ro04690Contr-R CTAGCGGGCCTCGTTCAGCCGTTC
16	Upstream region *fadE26*$_{DSM43269}$ (deletion construct)	1,234	fadE26UP-F TCGAGCAGTTCGGTAACGGTGAG fadE26UP-R AGATCTCGAGGTCCAGCGCGACATCATC
17	Downstream region *fadE26*$_{DSM43269}$ (deletion construct)	1,230	fadE26DOWN-F TTAGATCTTTCCTCGCGGAGGGCTTGTTGC fadE26DOWN-R GGCGAGATAGGCGACCAGATTC
18	*fadE26*$_{DSM43269}$ gene (confirmation Δ*fadE26* mutant)	520 (wt: 1,593)	fadE26Contr-F GCTTCACGCACGGCGTCCTCAC fadE26Contr-R GCGGTCCTCGGAGGCATCGAGT

Supplemental data for **Chapter 5**

Supplemental Table S1. List of primers used in this study. Restriction sites that were introduced are underlined.

PCR amplicon	size (bp)	oligonucleotide sequence
fragment *ltp3*$_{DSM43269}$ (screening genomic library	536	ltp3DEG-F GA(A/G)(A/T)(C/G)(A/C/G/T)AA(C/T)GC(A/C/G/T)A TGTGGGC
		ltp3DEG-R A(A/G)CCACAT(A/C/G/T)GG(C/T)TC(A/G)AACCA)
Upstream region *ltp3*$_{DSM43269}$ (deletion construct Δ*ltp3*Δ*ltp4*)	1,061	ltp3ltp4UP-F AAGTGCGGTCCGTCGCTGGTCA
		ltp3ltp4UP-R CATATGCGAGGCCGAGTGCCGGTAGTGC
Downstream region *ltp4*$_{DSM43269}$ (deletion construct Δ*ltp3* Δ*ltp4*)	1,052	ltp3ltp4DOWN-F CATATGGCGCTGTGGAAGCAGGCCGGTA
		ltp3ltp4DOWN-R AACTCGCGCAGGCGCGTGATGA
ltp3 ltp4$_{DSM43269}$ genes (confirmation Δ*ltp3* Δ*ltp4* deletion)	2,150 (wt: 4,525)	ltp3ltp4Contr-F GACCGCAGCGGATGCGTGATGG
		ltp3ltp4Contr-R CGCACACGGCCTGGATCTCGAA
Upstream region *ltp3*$_{DSM43269}$ (deletion construct Δ*ltp3*)	1,141	ltp3UP-F AGTCGTCGGCTTCGCGCATGCT
		ltp3UP-R CAGATCTGACGAGGCCGGACATCGACACG
Downstream region *ltp3*$_{DSM43269}$ (deletion construct Δ*ltp3*)	1,106	ltp3DOWN-F CAGATCTGGCGGTGTGCTCTGCTCCAATCC
		ltp3DOWN-R GCCGGATGTTCCTCGCTCATGC
ltp3$_{DSM43269}$ gene (confirmation Δ*ltp3* deletion)	1,214 (wt: 2,105)	ltpContr-F CGTGCTTCCAGGAGCTGTACGA
		ltpContr-R TACTGCGAACCACCGCCGTAGG
ltp3$_{DSM43269}$ gene (complementation Δ*ltp3*)	1,174	COMPltp3-F GGCGAAGCAACCCGCAGCTGTC
		COMPltp3-R GGTACCCTAGTTCGTGGGCCGGTCGGAG
Upstream region *ltp4*$_{DSM43269}$ (deletion construct Δ*ltp4*)	1,063	ltp4UP-F ACCGCTGAATCTGAAGTTCGAC
		ltp4UP-R GCATATGCTTGTCGTACAGCTCCTGGAAG
Upstream region *ltp4*$_{DSM43269}$ (deletion construct Δ*ltp4*)	1,064	ltp4DOWN-F GCATATGGCAGCGCAGGCGATCATGGACG
		ltp4DOWN-R GTCTCGCCGGCCTCGGTGAGTT
ltp4$_{DSM43269}$ gene (confirmation Δ*ltp4* deletion)	1,271 (wt: 2,105)	ltpContr-F CGTGCTTCCAGGAGCTGTACGA
		ltpContr-R TACTGCGAACCACCGCCGTAGG
ltp4$_{DSM43269}$ gene (complementation Δ*ltp4*)	1,066	COMPltp4-F GACAGACATCGCAGTCGTCGGCTTC
		COMPltp4-R GGTACCTCAGTTGCGTCCCTCCAGAACG
Internal fragment *fadA5*$_{DSM43269}$ (gene disruption)	734	fadA5-F GCCGTTGACGTTGACCTTGTCC
		fadA5-R CGCAGCAGGCCAACCACCTCAT
fadA5$_{DSM43269}$ (confirmation Ω*fadA5* disruption)	1,235	fadA5Contr-F GCCGTGCGACCAAGAGATCCG
		pK18Contr-R AATGCAGCTGGCACGACAGGTT